

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using bw model

Run on: March 23, 2006, 04:56:29 ; Search time 74.1768 Seconds
(without alignments)
1652.628 Million cell updates/sec

Title: US-10-617-038-21
Perfect score: 1410
Sequence: 1 MNOSHKPPSIYVIGIDGSKPA.....GNAVLQERAGCTLLVVGQOYL 279

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*
9: Geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1410	100.0	279	8	AD137300 M. tuberc
2	431.5	30.6	268	8	AD137321 M. tuberc
3	416	29.5	272	8	AD137306 M. tuberc
4	257	18.2	295	8	AD137298 M. tuberc
5	214	15.2	297	3	AAV87878 M. tuberc
6	214	15.2	297	5	ABG30907 Mycobacte
7	207	14.7	301	4	AAV71872 C. glutam
8	207	14.7	301	4	AAV71865 C. glutam
9	207	14.7	301	4	AAV71865 C. glutam
10	207	14.7	301	4	AAV71865 C. glutam
11	207	14.7	301	5	ABG80324 C. glutam
12	207	14.7	301	5	ABG80324 C. glutam
13	194	13.8	316	4	AAV62080 Propionib
14	194	13.8	316	6	ABM58599 Propionib
15	180	12.8	300	6	ABU25937 Protein e
16	127.5	9.0	3227	8	ADQ91710 Polypeptid
17	127.5	9.0	3227	8	ABE87002 Streptom
18	127.5	9.0	3227	8	ABE87002 Streptom
19	118.5	8.4	2539	6	ABU33731 Protein e
20	115	8.2	3562	4	AAE82213 Polypeptid
21	113	8.0	7068	4	AAE10142 Streptom
22	113	8.0	9477	4	AAE10144 Streptom
23	112	7.9	11096	4	AAE10129 Streptom
24	109.5	7.8	10917	6	AAE36132 Streptom

25	107.5	7.6	801	8	ADN26185	Adn26185 Bacterial
26	107	7.6	815	7	ABO78505	ABO78505 Pseudomon
27	106.5	7.6	860	9	ADY26642	ADY26642 Streptom
28	106.5	7.6	3546	4	AAE82212	AAE82212 Polypeptid
29	106.5	7.6	9510	6	AAE35119	AAE35119 Streptom
30	106	7.5	1641	4	ABG98879	ABG98879 S. cinna
31	106	7.5	3753	4	ABG98878	ABG98878 S. cinna
32	105.5	7.5	152	6	ABU18053	ABU18053 Protein e
33	105.5	7.5	426	7	ABO63304	ABO63304 Klebsiell
34	105.5	7.5	1976	9	ADY80057	ADY80057 Amino aci
35	105.5	7.5	6291	7	AD123927	AD123927 Streptom
36	104	7.4	154	5	ABE48055	ABE48055 Listeria
37	104	7.4	4032	8	ADV98987	ADV98987 Nanchangm
38	103	7.3	3025	4	ABG98868	ABG98868 S. cinna
39	103	7.3	5435	4	AAE10145	AAE10145 Streptom
40	103	7.3	7429	9	ADY80055	ADY80055 Amino aci
41	102.5	7.3	1657	8	ADQ91704	ADQ91704 Polypeptid
42	102.5	7.3	1657	8	ABE86996	ABE86996 Streptom
43	102.5	7.3	1657	8	ABE86996	ABE86996 Streptom
44	102.5	7.3	10625	9	ADZ51727	ADZ51727 PKS FecC
45	102	7.2	455	2	AAE35205	AAE35205 Lipase #2

ALIGNMENTS

RESULT 1
AD137300 standard; protein; 279 AA.
ID AD137300;
AC AD137300;
XX
XX
DT 22-APR-2004 (first entry)
DE M. tuberculosis low oxygen induced antigen Rv2028c SEQ ID NO:21.
KW mycobacterial infection; vaccine; tuberculosis;
KW Mycobacterium tuberculosis; immunisation; antibacterial; gene therapy;
KW low oxygen induced antigen.
XX
XX OS Mycobacterium tuberculosis.
XX MO2004006952-A2.
XX PN
XX PD 22-JAN-2004.
XX PF 08-JUL-2003; 2003MO-DK00477.
XX PR 13-JUL-2002; 2002DX-00001098.
XX PA (STRAT-) STATENS SERUM INST.
XX PT Andersen P, Rosenkrands I, Stryhn A,
XX WPI; 2004-122778/12.
XX DR N-PSDB; AD137345.
XX Use of one or more polypeptides or their fragments, which are expressed
XX during the latent stage of the mycobacterial infection, and/or nucleic
XX acids encoding the polypeptides, for a therapeutic vaccine against
XX tuberculosis.
XX PS Claim 3; SEQ ID NO 21; 76pp; English.
XX The present invention describes polypeptides or their fragments, which
XX are expressed during the latent stage of a mycobacterial infection,
XX and/or nucleic acids encoding the polypeptides, which are useful for
XX creating a therapeutic vaccine against tuberculosis. Also de-scribed: (1)
XX a therapeutic vaccine against tuberculosis comprising one or more
XX polypeptides; (2) a method for treating an animal, including a human
XX being, with tuberculosis caused by virulent mycobacteria, e.g. by
XX Mycobacterium tuberculosis, M. africanum or M. bovis; (3) a method for
XX immunising an animal, including a human being, against tuberculosis

CC caused by virulent mycobacteria; (4) a method of diagnosing tuberculosis
CC caused by virulent mycobacteria in an animal, including a human being;
CC (5) a method for diagnosing previous or ongoing infection with a virulent
CC mycobacterium; and (6) a method of diagnosing Mycobacterium tuberculosis
CC infection in a subject. The polypeptides have antibacterial activities,
CC and can be used in vaccines and in gene therapy. The polypeptides are
CC useful for the manufacture of a therapeutic vaccine for treating an
CC individual who is infected by a virulent mycobacterium, e.g. M.
CC tuberculosis, and who is not vaccinated with BCG against tuberculosis.
CC The present sequence represents a low oxygen induced antigen, which is
CC used in the exemplification of the present invention.

XX
SQ Sequence 279 AA:

Query Match 100.0%; Score 1410; DB 8; Length 279;
Best Local Similarity 100.0%; Pred. No. 6.3e-138;
Matches 279; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNQSHKRPSTIVGIDGSKPAVQALMVADEASRDIPRLIYALIEPDDPGYAAHGAARK 60
DB 1 MNQSHKRPSTIVGIDGSKPAVQALMVADEASRDIPRLIYALIEPDDPGYAAHGAARK 60
QY 61 LAAENAVRYAFTVAEADRPVKVEVEITQERPTSLIRASAAALVCVGAIGVHFRPE 120
DB 61 LAAENAVRYAFTVAEADRPVKVEVEITQERPTSLIRASAAALVCVGAIGVHFRPE 120
QY 121 RVGSTAAALALSAQCQPAIVRPHRVPIGRDAWIVVEADGSSDYGVLGAVMAEARLDS 180
DB 121 RVGSTAAALALSAQCQPAIVRPHRVPIGRDAWIVVEADGSSDYGVLGAVMAEARLDS 180
QY 181 PVRVVTGROSGVGTGDDVRASLDRWLARMQPRYPDVAVGSAAHGELLDYLAGLGRSVH 240
DB 181 PVRVVTGROSGVGTGDDVRASLDRWLARMQPRYPDVAVGSAAHGELLDYLAGLGRSVH 240
QY 241 MVLVTSASDQEHVEQLVGAPNAVLOEAGCTLLVVGQOYL 279
DB 241 MVLVTSASDQEHVEQLVGAPNAVLOEAGCTLLVVGQOYL 279

RESULT 2
ADI37321
ID ADI37321 standard; protein; 268 AA.

AC ADI37321;
DT 22-APR-2004 (first entry)
XX
DE M. tuberculosis low oxygen induced antigen Rv1334c SEQ ID NO:42.
XX
KM mycobacterial infection; vaccine; tuberculosis;
KM Mycobacterium tuberculosis; immunisation; antibacterial; gene therapy;
XX low oxygen induced antigen.
XX
OS Mycobacterium tuberculosis.
XX
PN WO2004006952-A2.
XX
PD 22-JAN-2004.
XX
PF 08-JUL-2003; 2003WO-DK000477.
XX
PR 13-JUL-2002; 2002DK-00001098.
XX
PA (STAT-) STARENS SERUM INST.
XX
PI Andersen P, Rosenkrands I, Stryhn A;
XX
XX WPI; 2004-122778/12.
XX
XX DR N-PSDB; ADI37366.
XX
XX
PT Use of one or more polypeptides or their fragments, which are expressed
PT during the latent stage of the mycobacterial infection, and/or nucleic
PT acids encoding the polypeptides, for a therapeutic vaccine against

PT tuberculosis.
XX
XX Claim 3; SEQ ID NO 42; 76pp; English.
XX
XX The present invention describes polypeptides or their fragments, which
XX are expressed during the latent stage of a mycobacterial infection,
XX and/or nucleic acids encoding the polypeptides, which are useful for
XX creating a therapeutic vaccine against tuberculosis. Also described: (1)
XX a therapeutic vaccine against tuberculosis comprising one or more
XX polypeptides; (2) a method for treating an animal, including a human
XX being, with tuberculosis caused by virulent mycobacteria, e.g. by
XX Mycobacterium tuberculosis, M. africanum or M. bovis; (3) a method for
XX immunising an animal, including a human being, against tuberculosis
XX caused by virulent mycobacteria; (4) a method of diagnosing tuberculosis
XX caused by virulent mycobacteria in an animal, including a human being;
XX (5) a method for diagnosing previous or ongoing infection with a virulent
XX mycobacterium; and (6) a method of diagnosing Mycobacterium tuberculosis
XX infection in a subject. The polypeptides have antibacterial activities,
XX and can be used in vaccines and in gene therapy. The polypeptides are
XX useful for the manufacture of a therapeutic vaccine for treating an
XX individual who is infected by a virulent mycobacterium, e.g. M.
XX tuberculosis, and who is not vaccinated with BCG against tuberculosis.
XX The present sequence represents a low oxygen induced antigen, which is
XX used in the exemplification of the present invention.

XX
SQ Sequence 268 AA:

Query Match 30.6%; Score 431.5; DB 8; Length 268;
Best Local Similarity 40.4%; Pred. No. 4.3e-36;
Matches 111; Conservative 34; Mismatches 105; Indels 25; Gaps 8;

QY 1 MNQSHKRPSTIVGIDGSKPAVQALMVADEASRDIPRLIYALIEPDDPGYAAHGAARK 60
DB 1 MNDPRPARAVVGVGIDGSKPAVQALMVADEAVRDIPLRLIYALIDPSQSAAGEGQ-- 58
QY 61 LAAENAVRYAFTVAEADRPVKVEVEITQERPTSLIRASAAALVCVGAIGVHFRPE 120
DB 59 -SARAAALHDASKKVEATGVPKIEVLCGRPLTKMQSSRAALCVSGVGLDVRGR 117
QY 121 RVGSTAAALALSAQCQPAIVRPH-RVPIGRDAWIVVEADGSSDYGVLGAVMAEARL 178
DB 118 R-GVAATLALGSLCQVAVIHPSRAPATTSQVSAVAEVDN---GVTLRHAFFEARLR 172
QY 179 DSPRVVTCROSGVGTGDDV-----RASLDRWLARMQPRYPDVAVGSAAHGELLDY 231
DB 173 GVPRLRAVAVH--AAETPDDEGSRRLAHVHLSRRLAHMTRLYRVEVRVRLAIGGSACRH 229
QY 232 LAGLGRSVHNVVLSADQEHVEQLVGA--PGNAVLT 264
DB 230 LAANAKPGQLFV--ADSHSAHELCGAYOPGCAVL 261

RESULT 3
ADI37306
ID ADI37306 standard; protein; 272 AA.

AC ADI37306;
DT 22-APR-2004 (first entry)
XX
DE M. tuberculosis low oxygen induced antigen Rv2624c SEQ ID NO:27.
XX
XX
KM mycobacterial infection; vaccine; tuberculosis;
KM Mycobacterium tuberculosis; immunisation; antibacterial; gene therapy;
XX low oxygen induced antigen.
XX
OS Mycobacterium tuberculosis.
XX
PN WO2004006952-A2.
XX
PD 22-JAN-2004.
XX
XX
XX 08-JUL-2003; 2003WO-DK000477.

XX 13-JUL-2002; 2002DK-00001098.
 PR (STAT-) STATENS SERUM INST.
 PA Andersen P, Rosenkrands I, Stryhn A;
 PI WPI; 2004-122778/12.
 DR N-PSDB; ADI37351.
 XX
 PT Use of one or more polypeptides or their fragments, which are expressed
 PT during the latent stage of the mycobacterial infection, and/or nucleic
 PT acids encoding the polypeptides, for a therapeutic vaccine against
 PT tuberculosis.
 PS
 PS Claim 3; SEQ ID NO 27; 76pp; English.
 XX
 CC The present invention describes polypeptides or their fragments, which
 CC are expressed during the latent stage of a mycobacterial infection,
 CC and/or nucleic acids encoding the polypeptides, which are useful for
 CC creating a therapeutic vaccine against tuberculosis. Also described: (1)
 CC a therapeutic vaccine against tuberculosis comprising one or more
 CC polypeptides; (2) a method for treating an animal, including a human
 CC being, with tuberculosis caused by virulent Mycobacterium, e.g. by
 CC Mycobacterium tuberculosis, M. africanum or M. bovis; (3) a method for
 CC immunising an animal, including a human being, against tuberculosis
 CC caused by virulent mycobacteria; (4) a method of diagnosing tuberculosis
 CC caused by virulent mycobacteria in an animal, including a human being;
 CC (5) a method for diagnosing previous or ongoing infection with a virulent
 CC mycobacterium; and (6) a method of diagnosing Mycobacterium tuberculosis
 CC infection in a subject. The polypeptides have antibacterial activities,
 CC and can be used in vaccines and in gene therapy. The polypeptides are
 CC useful for the manufacture of a therapeutic vaccine for treating an
 CC individual who is infected by a virulent mycobacterium, e.g. M.
 CC tuberculosis, and who is not vaccinated with BCG against tuberculosis.
 CC The present sequence represents a low oxygen induced antigen, which is
 CC used in the exemplification of the present invention.
 XX
 SO Sequence 272 AA;
 Query Match 29.5%; Score 416; DB 8; Length 272;
 Best Local Similarity 36.8%; Pred. No. 1.8e-34;
 Matches 98; Conservative 46; Mismatches 116; Indels 6; Gaps 2;
 QY 9 SIIVGIDSGKPAVOALMAVDEASRDIPRLVLAIEPDDPGVAHGAARKLAAENAV 68
 DB 11 TIVGIDSGHAAITPAALMGVDEATSRAPLVIVIRTPSPDYD--RDLHAERSL 67
 QY 69 RYAFATAEADRPVKEVEITQERPTSLIRASAAALVCGAIGVHFRPERVSTAA 128
 DB 68 RMAQSAVEAAGKLVKIFTDIPRGPAGPVIVASHADAMICGSGIGIRVASSILGSTATE 127
 QY 129 LAISAQCPVALIRPHRVPIGRDAAIYVEADGSSDIGVLGAVMAEARLDRSPVRYTCR 188
 DB 128 LAEKARCPVAVMRSKVDQPADINMIVMTDADNNAVLEYAAREAKLQAAPITALGGR 187
 QY 189 QSGVGIDGDDVRASLDLWFLARWQRPYRVQSAAYGELLDTAGAGRSYHMYVLSASD 246
 DB 188 PEBRRELPD--GEFERVDDMHRRHPDVRYPITTTGTARFLADHDERVQLAVIGGGE 244
 QY 249 QEHVEQLVGAAGNAVLOEAGCTLLV 274
 DB 245 AGQARLVGPSGHVFRHAECSVLV 270
 RESULT 4
 ADI37298
 ID ADI37298 standard; protein; 295 AA.
 XX
 AC ADI37298;
 XX
 DT 22-APR-2004 (first entry)

DE M. tuberculosis low oxygen induced antigen Rv2005c SEQ ID NO:19.
 XX mycobacterial infection; vaccine; tuberculosis;
 KM Mycobacterium tuberculosis; immunisation; antibacterial; gene therapy;
 KM low oxygen induced antigen.
 XX
 OS Mycobacterium tuberculosis.
 XX
 PN WO2004006952-A2.
 XX
 PD 22-JAN-2004.
 XX
 PF 08-JUL-2003; 2003WO-DK000477.
 XX
 PR 13-JUL-2002; 2002DK-00001098.
 XX
 PA (STAT-) STATENS SERUM INST.
 XX
 PI Andersen P, Rosenkrands I, Stryhn A;
 DR WPI; 2004-122778/12.
 DR N-PSDB; ADI37343.
 XX
 PT Use of one or more polypeptides or their fragments, which are expressed
 PT during the latent stage of the mycobacterial infection, and/or nucleic
 PT acids encoding the polypeptides, for a therapeutic vaccine against
 PT tuberculosis.
 PS
 PS Claim 3; SEQ ID NO 19; 76pp; English.
 XX
 CC The present invention describes polypeptides or their fragments, which
 CC are expressed during the latent stage of a mycobacterial infection,
 CC and/or nucleic acids encoding the polypeptides, which are useful for
 CC creating a therapeutic vaccine against tuberculosis. Also described: (1)
 CC a therapeutic vaccine against tuberculosis comprising one or more
 CC polypeptides; (2) a method for treating an animal, including a human
 CC being, with tuberculosis caused by virulent Mycobacterium, e.g. by
 CC Mycobacterium tuberculosis, M. africanum or M. bovis; (3) a method for
 CC immunising an animal, including a human being, against tuberculosis
 CC caused by virulent mycobacteria; (4) a method of diagnosing tuberculosis
 CC caused by virulent mycobacteria in an animal, including a human being;
 CC (5) a method for diagnosing previous or ongoing infection with a virulent
 CC mycobacterium; and (6) a method of diagnosing Mycobacterium tuberculosis
 CC infection in a subject. The polypeptides have antibacterial activities,
 CC and can be used in vaccines and in gene therapy. The polypeptides are
 CC useful for the manufacture of a therapeutic vaccine for treating an
 CC individual who is infected by a virulent mycobacterium, e.g. M.
 CC tuberculosis, and who is not vaccinated with BCG against tuberculosis.
 CC The present sequence represents a low oxygen induced antigen, which is
 CC used in the exemplification of the present invention.
 XX
 SO Sequence 295 AA;
 Query Match 18.2%; Score 257; DB 8; Length 295;
 Best Local Similarity 29.4%; Pred. No. 7.2e-18;
 Matches 91; Conservative 43; Mismatches 126; Indels 50; Gaps 9;
 QY 1 MNSHSPSTIVGIDSGKPAVOALMAVDEASRDIPRLVLAIEPDDPGVAHGAARKLAAENAV 47
 DB 1 MSKPRQHGIVGVVSDLSLSDAAACGATDAARNIPITVHVAVNADVATWPMYPTW 60
 QY 48 -----DPGYAAHGAARKLAAENAVRYAFATAEADRPVKEVEITQERPTSLIRASA 102
 DB 61 GVMQDEDE-----RQIYA--NAVKLAKEN--GADRKLSVSELVFTFPTWWEISN 109
 QY 103 AALVCGAIGVHFRPERVSTAAALALSAQCPVALIR-----PBRVP--IGRDA 152
 DB 110 EAEMVVLGSSGRGALRGLGVSSSLVRRACGPVAVIHSDAVIPDPQARPLVIGDGS 169
 QY 153 WI-----VVEADSSSDIGVLGAVMAEARLDRSPVRYRVTCRSGVGDDTGDVAAASLDRL 207
 DB 170 PVSELATVAVAFDEASRGVGLLAVHAW-----SDVEVELPGLDFAVQOGEALSLAERL 224

QY 208 ARWQPRYPDVVQSAAVHGEILDYLAGISVHMVVLASDQEHVE-QLVGAQENAVLQOE 266
 DB 225 AGMQERYPDVPSRVVVCDBRPARKLVQKSAQQLVVGSHRGGLTGMLIGSVSNVAVLHA 284
 QY 267 AGCTLLVVGQ 276
 DB 285 ARVPVIVARQ 294
 RESULT 5
 AAY87878
 ID AAY87878 standard; protein; 297 AA.
 XX
 AC AAY87878;
 XX
 DT 06-OCT-2000 (first entry)
 XX
 DE M. tuberculosis antigen TB32 protein.
 XX
 KM Tuberculosis; TB; antigen; vaccine; diagnosis; somatic; tuberculostatic;
 KM infection; interferon-gamma; IFN-gamma; protective immunity; therapy;
 KM delayed type hypersensitivity response; TB32.
 XX
 OS Mycobacterium tuberculosis.
 XX
 PN WO200021983-A2.
 PD 20-APR-2000.
 XX
 PF 08-OCT-1999; 99WO-DK000538.
 XX
 PR 08-OCT-1998; 98DK-00001281.
 PR 21-JAN-1999; 99US-0116673P.
 XX
 PA (STAT-) STATENS SERUM INST.
 XX
 PI Andersen P, Weidling K, Hansen CV, Florio W, Okkels LMM;
 PI Skjot RLV, Rosenkrande I;
 XX
 DR WPI; 2000-317931/27.
 DR N-PSDB; AAA39572.
 XX
 PT Novel polypeptide of somatic protein extract useful as vaccine against
 PT virulent Mycobacterium infection. Isolated from cell wall, cell membrane
 PT and cytosol.
 XX
 PS Claim 1; Page 106-107; 126pp; English.
 CC This invention describes a novel polypeptide (PP) of somatic proteins
 CC extract (I) which have tuberculostatic activity. (I) or their subsequence
 CC has at least one of the following properties: (a) the PP induces an in
 CC vitro recall response, or an in vitro response, during primary infection
 CC with virulent Mycobacterium, determined by a release of interferon (IFN)-
 CC gamma, (b) PP induces a protective immunity, determined by vaccinating an
 CC animal with PP and an adjuvant, three times at two weeks intervals, (c)
 CC PP induces an in vitro response, or in vitro recall response, determined
 CC by release of IFN-gamma of at least 1000 pg/ml or 500 pg/ml,
 CC respectively, from peripheral blood mononuclear cells (PBMC) withdrawn
 CC from TB patients, or PPD positive individuals, 6 months after diagnosis,
 CC (d) PP induces a specific antibody response in a TB patient, as
 CC determined by enzyme linked immunosorbent assay (ELISA) technique or a
 CC western blot, (e) PP induces a positive delayed type hypersensitivity
 CC (DTH) response, determined by intradermal injection. (I) and (II) are
 CC useful in preparing a prophylactic or therapeutic medicine as a vaccine
 CC for induction of a protective or generation of an immune response in a
 CC mammal against infection with a virulent Mycobacterium. (I) and (II) are
 CC also useful as diagnostic reagent for the diagnosis of a virulent
 CC Mycobacterium infection. The vaccine of the invention induces efficient
 CC immunological memory, providing long term protection against TB. This
 CC sequence represents a Microbacterium tuberculosis TB32 antigen described
 CC in the invention
 CC
 SQ Sequence 297 AA;

Query Match 15.2%; Score 214; DB 3; Length 297;
 Best Local Similarity 27.2%; Pred. No. 2.2e-13;
 Matches 84; Conservative 37; Mismatches 134; Indels 54; Gaps 10;
 QY 1 MNQSHKRPISIVGIDGSKPAVQALWAVDEASRDIPRLVLAIFPD-----DPQY 51
 DB 1 MSSGNSISLIGIVGIDSPAAQVAVRWAAADAEIRKIPTLVHAHSVPAVTWLEVPILPGV 60
 QY 52 AA---HGAARKLAENAVRYAFTVAEAD---RPVVEVEITQERPVTSILIRASAA 104
 DB 61 LRWQDGHG-----RHIDDLKVVQQLSLRACRPVYHSEIVPAAVPTLVDSKXA 111
 QY 105 ALVCVGAIGVHHRRPVRVSTAAALASACCPVALV-----RPHRVPIGRDAAMIV 156
 DB 112 VLMVVGCLSGRWPGRHLGVSISGLLRHAHCPVVIITHEDSVMPHFQQA-----VLV 164
 QY 157 EADGSSDIGVLGAVVAEALRDSPPRVVYTCRQSGGDT-----GDVVRAS-----LDR 205
 DB 165 GVDGSSASELATRIAFDEASRRN--VDLVALLHMSDVDSBWPGLDWPATQSMAEQVLA 222
 QY 206 WLRMQPRYPDVVQSAAVHGEILDYLAGISVHMVVLASDQ-EHVEQLVGAPGNVAVL 264
 DB 223 RLAGMQERYPNVAITRVVTRDQPARQLVQSEBAQLVVGSRGRGYACMLVGSVGETVA 282
 QY 265 QEAGCTLLV 273
 DB 283 QLARTPVIV 291
 RESULT 6
 ABG30907
 ID ABG30907 standard; protein; 297 AA.
 XX
 AC ABG30907;
 XX
 DT 21-OCT-2002 (first entry)
 XX
 DE Mycobacterium tuberculosis dormancy-induced protein Rv2623.
 XX
 KM Anti-mycobacterial; Mycobacterium tuberculosis; strain H37Rv;
 KM non-oxygen limiting; hypoxic stationary; hypoxic growth phase;
 KM tuberculosis; dormant mycobacterial infection; dormancy-induced; Rv2623;
 KM tuberculostatic; vaccine.
 XX
 OS Mycobacterium tuberculosis.
 XX
 PN WO200248391-A2.
 PD 20-JUN-2002.
 XX
 PF 11-DEC-2001; 2001WO-EP014551.
 XX
 PR 13-DEC-2000; 2000GB-00030368.
 XX
 PA (MOLE-) INST MOLECULAR & CELL BIOLOGY.
 XX
 PI Dick T, Calvin BKK;
 XX
 DR WPI; 2002-583530/62.
 DR N-PSDB; ABR9546.
 XX
 PT Identifying an anti-mycobacterial agent that modulates
 PT activity/expression of a protein expressed by Mycobacterium, involves
 PT monitoring the effect of an agent on the activity/expression of the
 PT protein or polynucleotide/vector encoding it.
 XX
 PS Example 2; Page 44-45; 56pp; English.
 CC The present invention relates to a new method for the identification of
 CC anti-mycobacterial agent that modulates activity and/or expression of
 CC protein expressed by Mycobacterium in non-oxygen limiting or hypoxic
 CC stationary, hypoxic growth phase. The method involves contacting a test

XX	Sequence	301 AA;	14.7%;	Score 207;	DB 4;	Length 301;
CC	Query Match		26.1%;	Pred. No. 1.2e-12;		
CC	Best Local Similarity		46;	Mismatches 134;	Indels 38;	Gaps 10
CC	Matches	77;	Conservative			
QY	10	IVVGIDGSKPAVOALMAVDEAASRDIPLRL-----LVA---IEPDDPGYAAHGA	57			
DB	7	VVVAADGSGASQAVRMAANTANKGIPRLIASSYTMPQFLYEGNVPQDELFDLQAE	66			
QY	58	ARKLAAENAVYVAFVAADRPVVEVEIQERVSTLIRSAALALVCVAGIGVHH	117			
DB	67	LEKINEAD-----IAHEVAD-IKIGHTIAGSTIDMLLEMSPDATYIVNGSRGLGL	119			
QY	118	RPERVGSATAALALSAQCPVAIVRPHRVVIGCDAAW--IVVEADGSSDILGVLLGA	175			
DB	120	SGMWGVSIGAVNVAHAKCPVVVVREBDA-VNDSKTKGPVVGVGDSSEVQQA	178			
QY	176	RLR-----DSPVRVVTGROSQVGDGTDDV-RASLDRLMAWQ--RYPRVQ	220			
DB	179	EARGAELVAVHTMTDMQVQASLAGLAAQQQWDEVERQOTDMLIERLAPLVEKYP	238			
QY	221	SAAVHGEILDYIAGLGRSVHMTVLSASDENE-QLVGA	274			
DB	239	KITTRDPRVRLAENAVNAQLLVGSHGRGFGKMLLGSTSRALLQSAPCPMMV	293			
RESULT 8						
AAU71872						
ID	AAU71872	standard; protein; 301 AA.				
XX	AAU71872;					
DT	26-FEB-2002	(first entry)				
XX						
DE	C. glutamicum metabolic pathway protein encoded by gene #7.					
XX						
KW	Metabolic pathway protein; MP; lysine biosynthesis pathway;					
KW	methionine biosynthesis pathway; large-scale production of fine chemical;					
XX	Corynebacterium diphtheriae; diphtheria.					
OS	Corynebacterium glutamicum.					
XX						
XX	MO20016573-A2.					
PD	13-SEP-2001.					
XX						
XX	22-DEC-2000; 2000MO-IB002035.					
XX						
XX	09-MAR-2000; 2000US-0187970P.					
XX	23-JUN-2000; 2000US-00606740.					
XX						
PA	(BADI) BASF AG.					
XX						
PI	Pompejus M., Kroeger B., Schroeder H., Zelder O., Habermann G;					
PI	Kim J., Lee H., Hwang B;					
XX						
DR	WPI. 2001-582369/65.					
XX	N-P8DB; AAS96082.					
PT						
PT	Nucleic acids encoding metabolic pathway proteins from Corynebacterium					
PT	glutamicum, useful for producing methionine and lysine in Corynebacterium					
XX	and Brevibacterium.					
XX						
PS	Disclosure: Page 169-170, 316pp; English.					
XX						
CC	The present invention relates to the isolation of novel Corynebacterium					
CC	glutamicum genes (AAS96073-AAS96132) encoding metabolic pathway (MP)					
CC	proteins. The metabolic pathway proteins of the invention include enzymes					
CC	involved in the lysine and methionine biosynthetic pathways. The					
CC	polynucleotide sequences of the invention can be used for the large-scale					
CC	production and/or modulation of expression of fine chemicals such as					

[illegible]

CC proteins. The metabolic pathway proteins of the invention include enzymes
 CC involved in the lysine and methionine biosynthetic pathways. The
 CC polynucleotide sequences of the invention can be used for the large-scale
 CC production and/or modulation of expression of fine chemicals such as
 CC lysine and methionine. The sequences of the invention may be used to
 CC identify C. glutamicum and related organisms e.g. C. diphtheriae in a
 CC subject to detect diphtheria. AAU71863-AAU71922 represent the novel C.
 CC glutamicum metabolic pathway proteins of the invention
 XX

SO Sequence 301 AA;

Query Match 14.7%; Score 207; DB 4; Length 301;
 Best Local Similarity 26.1%; Pred. No. 1.2e-12;

Matches 77; Conservative 46; Mismatches 134; Indels 38; Gaps 10;

```

QY 10 IVGGIDSKPAVQALMAVDEAASRDIPRL-----LYA---IEPDDPGYAAGAA 57
D 7 VVAVVDSDASKAVRMAANTANKRGIPRLASSTYMPQFLYAGMTVPQELFPDLOAEA 66
QY 58 ARKLAAMAVRYAFTAVEAADRPVKEVEITQERPTSLIRASAAALVCVGAIGVHHF 117
D 67 LEKINEARD-----IAHEVAPE-IKIGHTIAEGSPIDMLLEMSPDATMIWGSRGIGL 119
QY 118 RPEVGSSTAALALSAQCPVALVRPHRVPFGDAA--IYVEADGSSDIDGLGAWMAEA 175
D 120 SGWMSGVSGAVVSHAKCPVVVVEDSA-VNEDSKYGPVVVGVDGSEVSOQATEYAFAEA 178
QY 176 RLR-----DSPRVVTCRQSGVGTGDDV-RASLDRLMARQP---RYPDVAVQ 220
D 179 EARGAEIVAVHTMDVQASLAGLAAAOQOMDEVERQOTDMLIERLAPVEKIPSTYVK 238
QY 221 SAAVHGELIDYLAGISRVHVVLSASDOEHVE-OLVGAPGNALVQAGCTLLVV 274
D 239 KIIRDRPVRALAEASNAQLLVVSGHGRGFKMGLGSTRLAQSAFCMMV 293

```

RESULT 10
 AAG92991

ID AAG92991 standard; protein; 301 AA.

XX AAG92991;

XX 26-SEP-2001 (first entry)

XX C glutamicum protein fragment SEQ ID NO: 6745.

XX Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
 KW organic acid synthesis.

XX Corynebacterium glutamicum.

XX EPI108790-A2.

XX 20-JUN-2001.

XX 18-DEC-2000; 2000EP-00127688.

XX 16-DEC-1999; 99UP-00377484.

XX 07-APR-2000; 2000JP-00159162.

XX 03-AUG-2000; 2000JP-00280988.

XX (KYOWA) KYOWA HAKKO KOGYO KK.

XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
 PI Tateishi N, Senoh A, Ikeda M, Ozaki A;

XX MPI; 2001-376931/40.

XX N-PSDB; AAH68210.

XX Novel polynucleotides derived from Coryneform bacteria, for identifying
 PT mutation point of a gene, measuring expression of a gene, analyzing
 PT expression profile or pattern of a gene and identifying homologous gene.
 XX

PS Claim 17, SEQ ID NO 6745; 246bp + Sequence Listing; English.

XX The present invention provides a number of nucleotide and protein
 CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
 CC are useful for identifying the mutation point of a gene derived from a
 CC mutant of coryneform bacterium, measuring expression amount and analysing
 CC the expression profile or expression pattern of a gene derived from
 CC Coryneform bacterium, and identifying a homologue of a gene derived from
 CC coryneform bacterium. Coryneform bacteria are useful for producing amino
 CC acids, nucleic acids, vitamins, saccharides and organic acids,
 CC particularly L-lysine. The present sequence is a protein described in the
 CC exemplification of the invention. Note: The sequence data for this patent
 CC did not form part of the printed specification, but was obtained in
 CC electronic format directly from the European Patent Office
 XX

SO Sequence 301 AA;

Query Match 14.7%; Score 207; DB 4; Length 301;
 Best Local Similarity 26.1%; Pred. No. 1.2e-12;

Matches 77; Conservative 46; Mismatches 134; Indels 38; Gaps 10;

```

QY 10 IVGGIDSKPAVQALMAVDEAASRDIPRL-----LYA---IEPDDPGYAAGAA 57
D 7 VVAVVDSDASKAVRMAANTANKRGIPRLASSTYMPQFLYAGMTVPQELFPDLOAEA 66
QY 58 ARKLAAMAVRYAFTAVEAADRPVKEVEITQERPTSLIRASAAALVCVGAIGVHHF 117
D 67 LEKINEARD-----IAHEVAPE-IKIGHTIAEGSPIDMLLEMSPDATMIWGSRGIGL 119
QY 118 RPEVGSSTAALALSAQCPVALVRPHRVPFGDAA--IYVEADGSSDIDGLGAWMAEA 175
D 120 SGWMSGVSGAVVSHAKCPVVVVEDSA-VNEDSKYGPVVVGVDGSEVSOQATEYAFAEA 178
QY 176 RLR-----DSPRVVTCRQSGVGTGDDV-RASLDRLMARQP---RYPDVAVQ 220
D 179 EARGAEIVAVHTMDVQASLAGLAAAOQOMDEVERQOTDMLIERLAPVEKIPSTYVK 238
QY 221 SAAVHGELIDYLAGISRVHVVLSASDOEHVE-OLVGAPGNALVQAGCTLLVV 274
D 239 KIIRDRPVRALAEASNAQLLVVSGHGRGFKMGLGSTRLAQSAFCMMV 293

```

RESULT 11
 ABG80324

ID ABG80324 standard; protein; 301 AA.

XX ABG80324;

XX 15-NOV-2002 (first entry)

XX C. glutamicum metabolic pathway (MP) protein #4.

XX Metabolic pathway protein; MP; cell metabolism; amino acid; vitamin;
 KW cofactor; nucleotide; nucleoside; trehalose; fine chemical production;
 KW organic acid; non-proteinogenic amino acid; purine base; carbohydrate;
 KW pyrimidine base; lipid; unsaturated fatty acid; diol; polyketide;
 KW aromatic compound; food industry; animal feed; cosmetic industry;
 KW pharmaceutical industry; enzyme.

XX Corynebacterium glutamicum ATCC 13032.

XX MO200251231-A1.

XX 04-JUL-2002.

XX 22-DEC-2000; 2000WO-EP013143.

XX 22-DEC-2000; 2000WO-EP013143.

XX (BADI) BASF AG.

XX Pompejus M, Kroegeer B, Zelder O, Schroeder H;
 PI

AC AAU62080;
XX
DT 27-FEB-2002 (first entry)
XX
DE Propionibacterium acnes immunogenic protein #22976.
XX
XX
XX SAPHO syndrome; synovitis; acne; pustulosis; hyperostosis; osteomyelitis;
KM uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KM inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KM dermatological; osteoprotectic; neuroprotectant.
XX
OS Propionibacterium acnes.
XX
XX WO200181581-A2.
XX
XX
XX 01-NOV-2001.
XX
XX 20-APR-2001; 2001WO-US012865.
XX
XX 21-APR-2000; 2000US-0199047P.
PR 02-JUN-2000; 2000US-0208841P.
PR 07-JUL-2000; 2000US-0216747P.
XX
XX (CORI-) CORIXA CORP.
XX
XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
XX WPI; 2001-616774/71.
DR N-PSDB; AAS59623.
XX
XX
XX Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris.
XX
XX
XX Example 1; SEQ ID NO 23275; 1069pp; English.
XX
XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX
SQ Sequence 316 AA;
Query Match 13.8%; Score 194; DB 4; Length 316;
Best Local Similarity 25.2%; Pred. No. 2.8e-11;
Matches 68; Conservative 43; Mismatches 15; Indels 8; Gaps 3;

Db 165 TSVAVAGMSAPVAVVTSRASTPPTGKKCAVAAGVQSGVTAAGVGFGEADHRGKLT 224
Qy 184 VVTCROSGVDTGGDDVNASLDRMLARWQPRYPDVAVGSAVHGELPYLAGLGRSVAMV 243
Db 225 AVTVPGNDSEVHDEALKRLNEVVKPLANKPHVEVETRVLSGEPDALVDLSGVLLV 284
Qy 244 LSASDQHEVQLVGPAGNAVLQEAQCTLLV 273
Db 285 IGKKHPILGWTAGVGRALIMAHQSPAL 314
RESULT 14
ABMS58599
ID ABMS58599 standard; protein; 316 AA.
XX
XX
XX ABMS58599;
XX
XX 20-OCT-2003 (first entry)
XX
XX
XX Propionibacterium acnes predicted ORF-encoded polypeptide #23275.
DE
XX
XX Acne vulgaris; antiseborrhoeic; dermatological; antibacterial;
KM Immunostimulant; Immune response; vaccine.
XX
OS Propionibacterium acnes.
XX
XX WO2003033515-A1.
XX
XX 24-APR-2003.
XX
XX 11-OCT-2002; 2002MO-US032727.
XX
XX 15-OCT-2001; 2001US-00978825.
XX
XX
XX (CORI-) CORIXA CORP.
XX
XX Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;
PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;
PI Barth B, Vallieve-Douglas J;
XX
XX WPI; 2003-381789/36.
DR N-PSDB; ACF64552.
XX
XX
XX New Propionibacterium acnes polypeptides and polynucleotides encoding the
PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
PT or for stimulating an immune response specific for a P. acnes protein.
XX
XX Example 1; SEQ ID NO 23275; 1481pp; English.
XX
XX The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
CC encoding a Propionibacterium acnes protein. The invention also relates to
CC polynucleotides encoded by the polynucleotides (ABM35624-ABM64536) and to
CC immunogenic fragments of P. acnes polypeptides. The invention
CC additionally encompasses expression vectors and host cells comprising a
CC polynucleotide of the invention; antibodies against polypeptides of the
CC invention; fusion proteins comprising a polypeptide of the invention; a
CC method for stimulating an immune response specific for a P. acnes
CC polypeptide and an isolated T cell population comprising T cells prepared
CC via this method; a vaccine composition (comprising P. acnes polypeptides,
CC polynucleotides, antibodies, fusion proteins, T cell populations, or
CC antigen-presenting cells that express the polypeptide); a method and kit
CC for detecting or determining the presence or absence of P. acnes in a
CC patient; and a method for inhibiting the development of P. acnes in a
CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
CC proteins, T cell populations or antigen-presenting cells that express the
CC polypeptides are useful for diagnosing, preventing or treating acne
CC vulgaris, or for stimulating an immune response specific for a P. acnes
CC protein. The polynucleotides can also be used as probes or primers for
CC nucleic acid hybridisation. The vaccine composition is useful for the
CC stimulation of an immune response against P. acnes, or for treating acne,
CC and the kit is useful for performing a diagnostic assay. The present
CC sequence represents a polypeptide predicted to be encoded by an ORF (open

CC reading frame) contained within the P. acnes polynucleotides of the
 CC invention. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX

SQ Sequence 316 AA;

Query Match 13.8%; Score 194; DB 6; Length 316;
 Best Local Similarity 25.2%; Pred. No. 2.8e-11;
 Matches 68; Conservative 43; Mismatches 151; Indels 8; Gaps 3;

QY 10 IVVGIDSKPAVQALMAVDEAASRDIPRLVYIEPDDPGYAHGAARLTAENAVR 69
 DB 47 IIVGVDSSEGLRARAAYAGSAIRDDLLVLAHV--DDAAVAGAWGVYDPTLQDAQ 104
 QY 70 YA-----FTAVEADRPVKVEVEITQERPVTSILRASAAALVCVGAIGVHFRPERVGS 124
 DB 105 VVVDDALHVAIERMDPRISGEVVLGNPAAILDRSHDQLVLGRRATSGLERMVG 164
 QY 125 TAAALALSAQCPVAIVRPHRVPIGRDAAMIYVEADGSSDIG-VLLGAVMAEARLDSPPV 183
 DB 165 TSVAVAGMSAARVVVISRASTPDPDGKCVAVAVGPOSVGTAAVGFAEADHRGCKL 224
 QY 184 VVTCRQSGVGTGDDVRAASLDRLARKQRPYDVRVQSAVHGEILDTYLAGRGSYMMV 243
 DB 225 AVTVPGNDSEVHDEALKRLNEVVKPLANKPHVEVEITRVLGSPVALVDLSGVNLLV 284
 QY 244 LSASDQEHVEQLVGAPGNAVLOEAGCTLLV 273
 DB 285 IGMKHPILGWTAGVSRALMAHQSPLAT 314

RESULT 15

ABU25937
 ID ABU25937 standard; protein; 300 AA.

XX AC ABU25937;

XX DT 19-JUN-2003 (first entry)

XX DE Protein encoded by Prokaryotic essential gene #11464.

XX KW Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX OS Corynebacterium diphtheriae.

XX PN WO200277183-A2.

XX PD 03-OCT-2002.

XX PF 21-MAR-2002; 2002MO-US009107.

XX PR 21-MAR-2001; 2001US-00815242.

XX PR 06-SEP-2001; 2001US-00948893.

XX PR 25-OCT-2001; 2001US-0342923P.

XX PR 08-FEB-2002; 2002US-00072851.

XX PR 06-MAR-2002; 2002US-0362699P.

XX PA (ELIT-) ELITRA PHARM INC.

XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

XX PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX DR WPI; 2003-029926/02.

XX DR N-PSDB; ACA29807.

XX PT New antisense nucleic acids, useful for identifying proteins or screening

XX PT for homologous nucleic acids required for cellular proliferation to

XX PT isolate candidate molecules for rational drug discovery programs.

XX PS Claim 25; SEQ ID NO 53861; 1766bp; English.

XX CC The invention relates to an isolated nucleic acid comprising any one of

CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX

SQ Sequence 300 AA;

Query Match 12.8%; Score 180; DB 6; Length 300;
 Best Local Similarity 24.8%; Pred. No. 7.6e-10;
 Matches 73; Conservative 46; Mismatches 139; Indels 36; Gaps 9;

QY 10 IVVGIDSKPAVQALMAVDEAASRDIPRL-----LVA---IEPDDPGYAHGA 57
 DB 7 VVAVVDSSEASQNAVWMAANTANKRGVPLRLAASYTMPQFLVAGVWPQDELDELQSET 66
 QY 58 ARCLAAENAVRAFAVVEADRPVKVEVEITQERPVTSILRASAAALVCVGAIGVHNF 117
 DB 67 MDMEIAR-----VVAHEVAP-DIKIGYVLAGSPIDMLDMSDVMTIMVSGRGLGGL 119
 QY 118 RPERVGSTAALALSAQCPVAIVR-PHRVPIGRDAAMIYVEADGSSDIGVLLGAVMAEAR 176
 DB 120 SGWVMSVSAVSHDCFPVVVRSNDHTETNKGPVVVGVDGVSQGRATEFAEEBAQ 179
 QY 177 LRDSP-VRVVTCRQSGVGTGDDVRAASLDRL-----LARNQP---RYDVRVQGS 221
 DB 180 ARGAKLVAIHTWMDQVQASLAGLAAAOBWEIIEKEQTLLKDRILQPLERPPVEVEM 239
 QY 222 AAHGEILDYLAGRGSYHNVVTLASDQEHVE-QLVGAGGNNAVLOEAGCTLLV 274
 DB 240 VITDRPRVALDECANNAQLLVGSHGSGFRGMLIGSTRALLQSAFCMVVV 293

Search completed: March 23, 2006, 05:10:40
 Job time : 77.1768 sec

	Query Match	30.6%	Score 431.5;	DB 2;	Length 268;
	Best Local Similarity	40.4%;	Pred. No. 4.3e-26;		
	Matches 111;	Conservative 34;	Mismatches 105;	Indels 25;	Gaps 8;
Qy	1	MNQHSPESIVVIGIDGSKPAVOAALMAVDEASRDIPRLLYALEPDDPGYAAHGAARK	60		
Dd	1	MSDPRAPAVVVVGIDGSRRAATHAALMAVDENVNRDIPRLTYVVIDPSQLSAGGGGQ--	58		
Qy	61	LAAENAVRYFTAVEAADRPVKVEVEITQERPYSLSIRASAAAALVCVAIGVHHRRPE	120		
Dd	59	-SAARAAAHDSRKKEATGCPVKIETEVLCGRPLTKMQEBSRSAAMCVCVSGVDHVRGR	117		
Qy	121	RVGSTAAALALSAOCFPAIVRPH--RVPIGRDAAMIIVEALGSSSDIGVLCAWAAEARLR	178		
Dd	118	R-GSVAAATLASSALCPNAVIHPSRAEPRTTQGVSAVAEVDN---GVTLRHAEERLARLR	172		
Qy	179	DSPVKVVYTCROSGVGDTGDV-----RASLDRLMLAQRPDPDVAVQSAAVHGELLDY	231		
Dd	173	GVPLRAVAVH---AAETPDVVEGSGRLAHVHLSRLRAHWTRLYPEVVRDALIAGGSACRH	229		
Qy	232	LAGLGRSVHMVYLASDQEHVEQLVGA--PGNAVL	264		
Dd	230	LAANAKPGOLFV---ADSHSAHELCGAYOPCAVL	261		

Query Match 29.5%; Score 416; DB 2; Length 272;
Best Local Similarity 36.8%; Pred. No. 6.8e-25;
Matches 98; Conservative 46; Mismatches 116; Indels 6; Gaps 2;

QY 69 RYAF TAVEAD RPKVEVEITQERPVTS LIRASAAALVCVGAIGVHHFRPERVGSTAA 128

RESULT 4
 C70759
 hypothetical protein RV2005C - Mycobacterium tuberculosis (strain H37RV)
 C1Species: Mycobacterium tuberculosis
 C1Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
 C1Accession: C70759
 R1Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Felkewell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rastandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
 Nature 333, 537-544, 1998
 A1Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrall, B.G.
 A1Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
 A1Reference number: A70500; MUID:98295987; PMID:9634230
 A1Accession: C70759
 A1Status: preliminary; nucleic acid sequence not shown; translation not shown
 A1Molecule type: DNA
 A1Residues: 1295 <CO2>
 A1Cross-references: UNIPROT:Q10851; UNIPARC:UP1000013B62F; GB:274025; GB:AL123456; NID:93
 A1Experimental source: strain H37RV
 C1Genetics:
 A1Gene: RV2005C

Db 285 ARVPIVARQ 294

C;Accession: T36949

[illegible]

RESULT 8
T16963
hypothetical protein SCU_30c - Streptomyces coelicolor
C1Species: Streptomyces coelicolor
C1Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C1Accession: T16963
R1Seeger: K.J.; Harris, D.; Thomson, N.R.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, September 1999

A;Reference number: Z21607
A;Accession: T36963
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-328 <SEE>
A;Cross-references: UNIPROT:Q9RIY4; UNIPARC:UPI00000DB39C; EMBL:AL109962; PDB: CMB5148.
A;Experimental source: strain A3(2)
C;Genetics:
A;Gene: SCOEDB:SCJ1.30C

Query Match 15.0%; Score 211; DB 2; Length 328;
Best Local Similarity 29.3%; Pred. No. 5.5e-09;
Matches 91; Conservative 36; Mismatches 122; Indels 62; Gaps 12;

```
QY 9 SIIVGIDGSKPAVQALMVADEAASRDIPRLLYAIEBDDGVAHGAARKLAASNAV 68
D 32 TIVGIDGSPESHAAAEWAAAREALRDLVRLHWMED-----APAVLAQDSI 79
QY 69 RYATTAVEADNR-PVKV-----EVEITQER---PVTSLIRASAAAALVCVGAIGV 114
D 80 LGACTHOMHWERVQKVEGLRLHHPGVVVTSDORSQGPADTLVRDADGAEVLVLSRAP 139
QY 115 HHFRPERVGSFAALALSAOCPAIVR-----PHRY-----PIGRDAMIIVEAD-G 160
D 140 SGLGGLAGSVGOSVIHSETPVVLVRAGEQAGEHVDVSGVPPAANRRRPVVVGLDIG 199
QY 161 SSDIGVLGAVMABARLRDSPVRV-----TCRQSGVGDPTGDDVR--ASLDRWLA 208
D 200 SPDDGVLSFA-FDEAQGRGTAVHVGWGRPPYPSLTLAGVPPQGVARRKADLTLL 258
QY 209 RWCPRYDVVQSAAVHGGELLDYLAGRSVHNVVLSASDOE-----HVEQLVCAFGNAV 263
D 259 PMRSYSDVDEVEASRPGSPADLLAEASHNASLVVGRIRIRPPLGVH---IGAVNAHV 314
QY 264 LQEGAGCTLLV 274
D 315 LHRVAPVAIV 325
```

Db

RESULT 9
T36962
hypothetical protein SCJ1.29C - Streptomyces coelicolor

C;Species: Streptomyces coelicolor
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: T36962
R;Seeger, K.U.; Harris, D.; Thomson, N.R.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, September 1999
A;Reference number: Z21607
A;Accession: T36962
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-283 <SEE>
A;Cross-references: UNIPROT:Q9RIY5; UNIPARC:UPI00000DB39B; EMBL:AL109962; PDB: CMB5147.
A;Experimental source: strain A3(2)
C;Genetics:
A;Gene: SCOEDB:SCJ1.29C

Query Match 14.7%; Score 207; DB 2; Length 283;
Best Local Similarity 25.3%; Pred. No. 9.5e-09;
Matches 76; Conservative 49; Mismatches 121; Indels 54; Gaps 9;

```
QY 10 IVVGDGSKPAVQALMVADEAASRDIPRLLY-----AIEPDDPGVAHGAARK 60
D 5 LVVGVDSSESLRAVDAADEALHNAVPLWVFGDLMEYEGALAREPGKPTDMQADD 64
QY 61 LAAAEAAVRYAFTVAEADR--PVKEVEITQERPTVSLIRASAAAALVCVGAIGHFR 118
D 65 ILAA-----AAIRAGRHPDLVVTETVPDEAEHALICAGRNASMIWGSRGSGIA 116
QY 119 PEAVGSTAALALSAOCPAIVR---PHRVPIGRDAMIIVEADGSSDIGVLGAWAA 175
D 117 DRLLGSVSRIVAAAGSDCPVVVLKGNHNDRAIGKNGRIIVGVGVSAV---LRLAFMEA 173
```

```
QY 176 RLKDSPIRVVTCRQSGVGDPTGDDVRASLDRLAKMOP--RY-----PDV 217
D 174 RLKQVPAARVAARWCRPAHET-----IDHPLPAGEPARRYEERAAAREALEADAPDV 226
QY 218 RVQSAVHGGELLDYLAGRSVHNVVLSASDOEHVEQLVGAPEGNAVLOEAGCTLLVGOQ 277
D 227 AVRRYVEGPARAVLPAASABAGLLVIGRTQGR---LGRVAAVHLHRSACVTVVPPER 282
```

RESULT 10
B70758
hypothetical protein Rv1996 - Mycobacterium tuberculosis (strain H37RV)

C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C;Accession: B70758
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Comor, R.; Davies, R.; Devlin, K.; Feldwell, T.; Gentles, S.; Hamlin, N.; Holtroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A;Reference number: A70500; PMID:98295987; PMID:9634230
A;Accession: B70758
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-317 <COL>
A;Cross-references: UNIPROT:Q10862; UNIPARC:UPI00013B4FB; GB:274025; GB:AL123456; NID:G;
A;Experimental source: strain H37RV
C;Genetics:
A;Gene: Rv1996

Query Match 12.9%; Score 182.5; DB 2; Length 317;
Best Local Similarity 25.8%; Pred. No. 8.3e-07;
Matches 78; Conservative 43; Mismatches 138; Indels 43; Gaps 8;

```
QY 10 IVVGDGSKPAVQALMVADEAASRDIPRLLYAIEP---DDEGVA-----AHGAARK 61
D 10 IVVGVDSPCSHTAIVMAARDAQMRVALRVQVVPVITAPBGMFEYSRQEAARKRI 69
QY 62 AAENAVRYAFTVAEAA-----DRPVKEVEITQERPTVSLIRASAAAALVCVGA 111
D 70 VHSYLVQAQHQVDEQAHKVALEASSGGAQQTGTVLHGQIVPTLANISROVAMVLTG 129
QY 112 IGVHFRPERVGSSTAALALSAOCPAIV---RPHRVPIGRDAMIIVEADGSSDIGVL 167
D 130 RQGAVALGALIGSVSSLVHHAHPVAIVPEBRPARP---HAPVVGIDGSPFSGLA 185
QY 168 LGAVMAEARLRDSPVRVTCRQSGVG-----DTGDDVRSALDRMLAKMP 212
D 186 AEIAFDEASRRGVDL--VALHAWSDMGRLDFPRLNMAPIEMRLDEQEKMLARRLSGWOD 244
QY 213 RYVDVAVQSAAVHGGELLDYLAGRSVHNVVLSASDOEHVEQLVGAPEGNAVLOEAGCTL 271
D 245 RYVDVVHAKVVCDRAPRLLELAQTALQVLVVGSHRGGFPGMHLGVSRAVNVSGQAV 304
QY 272 LV 273
D 305 IV 306
```

Db

RESULT 11

T37031
hypothetical protein SCJ12.12C - Streptomyces coelicolor

C;Species: Streptomyces coelicolor
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: T37031
R;Murphy, L.; Harris, D.; Thomson, N.R.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, August 1999
A;Reference number: Z21619
A;Accession: T37031
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA

A;Residues: 1-301 <MUR>
A;Cross-references: UNIPROT:Q9RI46; UNIPARC:UPI00000DB3E6; EMBL:AL109989; PIDN:CAB53424.
A;Experimental source: strain A3(2)
C;Genetics:
A;Gene: SCOEDB:SCJ12.12C

Query Match 12.8%; Score 180.5; DB 2; Length 301;
Best Local Similarity 25.4%; Pred. No. 1.1e-06;
Matches 78; Conservative 40; Mismatches 136; Indels 53; Gaps 10;

QY 9 SIVVIGDSKPAVOALMAVDEASRDIPRLLYAIEPDDGYAAGAAAKLAABNAV 68
D 4 TITGLGSPSSRAAEWAAREGTLRRVRLHVWQP-----VPEPWAQAPLGCALRHQ 58
QY 69 RYAFVAEADRAKPKVE---VEITQER---PTSLIRASAAALVCGALGVHFRPER 121
D 59 HMTKRIPRDTAEGLRLHPGVETTEQATGNPADALLAGTIDABILVGSALSGLGF 118
QY 122 VGSTAAALASACCPVALVR-----PFRVPIGRDAW-----IIVEADGSSDYG 165
D 119 VGSVGVGVATERTPVLIVRAGEQADDEHLKDPGIP---SAATGFRVTVGLDTGSPDE 175
QY 166 VLLGAAVMAEARLDRSPRVVTCRQSGVGT-----GDDVRASLDRLMAR-----WQP 212
D 176 AVLSFAFEAEARRRRAPLTAIVAWNLPSSTYTSIAAGFPREELARAQALGELLPWRE 235
QY 213 RYPPVRVQSAVHGEILDYLAGLSRVHVVYLSASDQE-----HYEQLVGARGNAVLOEA 267
D 236 KYPPEVETETCRGLSPAEHLIDAAADASLVVGRIRKSPGCVH---TGAIVAHVMMHA 291
QY 268 GCTLLVV 274
D 292 TTPVAVV 298

RESULT 12

T37029
hypothetical protein SCJ12.10C - Streptomyces coelicolor
C;Species: Streptomyces coelicolor
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: T37029
R;Murphy, L.; Harris, D.; Thomson, N.R.; Parhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, August 1999
A;Reference number: Z21619
A;Accession: T37029
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-288 <MUR>
A;Cross-references: UNIPROT:Q9RI48; UNIPARC:UPI00000DB3E4; EMBL:AL109989; PIDN:CAB53422.
A;Experimental source: strain A3(2)
C;Genetics:
A;Gene: SCOEDB:SCJ12.10C

Query Match 12.1%; Score 170.5; DB 2; Length 288;
Best Local Similarity 25.2%; Pred. No. 6.3e-06;
Matches 76; Conservative 42; Mismatches 125; Indels 59; Gaps 11;

QY 10 IVVIGDSKPAVOALMAVDEASRDIPRLLYAIEPDDGYAAGAA-ARKL-----A 62
D 5 MVVGVDSSESLGAVDMADEALHEVPLRIVHAWRMD---RYEGASLARELKPQGHV 60
QY 63 AAEAAVRAFTAVEAADRPVAVVEITQERPVTSILIRASAAALVCGALGVHFRPERV 122
D 61 TTDTLAVATRRARRHHPDLAVTTTEATIEEPYVILIRARASAVIILSTRGRGELAGLL 120
QY 123 GSTAAALASACCPVALVR-----RVPIG-RDAAVIVE--ADGSSDYGVL 167
D 121 GSVSLVATMTSDCPVVTVRGSHDDRAAGGRGRIVGVADAPFAAVAFACEBARRGAA 180
QY 168 LGAVMA-----EAKLDSFVRVYTCR-----QSGVGTGDDVRASLDRLMARQP 212
D 181 LDAVAVRACPTHTDVHPLTAGTERLHEERAKLEBALADAPADV-----LRRRTA 234

QY 213 RYPPVRVQSAVHGEILDYLAGLSRVHVVYLSASDQEHVEQLVGAFGNAVLOEAGCTLL 272
D 235 EGPGRSVLSAASH-----EADLLVGRRRRPGFGHRLGRVAHTLLHRSACPVA 282

QY 273 VV 274
D 283 VV 284

RESULT 13

T36954
hypothetical protein SCJ1.21 - Streptomyces coelicolor
C;Species: Streptomyces coelicolor
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: T36954
R;Seeger, K.J.; Harris, D.; Thomson, N.R.; Parhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, September 1999
A;Reference number: Z21607
A;Accession: T36954
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-152 <SEE>
A;Cross-references: UNIPROT:Q9RI23; UNIPARC:UPI00000DB393; EMBL:AL109962; PIDN:CAB53139.
A;Experimental source: strain A3(2)
C;Genetics:
A;Gene: SCOEDB:SCJ1.21
C;Superfamily: Escherichia coli ybdQ protein

Query Match 9.6%; Score 136; DB 2; Length 152;
Best Local Similarity 29.8%; Pred. No. 0.0014;
Matches 45; Conservative 17; Mismatches 71; Indels 18; Gaps 2;

QY 1 MNOSHKRPISIVGIGDSKPAVOALMA-----VDEASRDIPRLLYAIEPDDPG 50
D 1 MQSEPTARVVGVDSPPSYAALRMADRYARAVGVVEAVHVDTSANGFRAPADPD 60
QY 51 YAAHGAARKLAAENAVRYAFTAVEAADRPVKEVEITQERPVTSILIRASAAALVCGV 110
D 61 FDLQARERFAELE-----ATFGERPPGKKEILVGDPEETLIRASGAGELLVVG 112
QY 111 ALGVHFRPERVVGSTAAALASACCPVALVR 141
D 113 RRGGAFAARMLGVSQRCQAQHAACPVVVR 143

RESULT 14

T36960
hypothetical protein SCJ1.27 - Streptomyces coelicolor
C;Species: Streptomyces coelicolor
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: T36960
R;Seeger, K.J.; Harris, D.; Thomson, N.R.; Parhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, September 1999
A;Reference number: Z21607
A;Accession: T36960
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-148 <SEE>
A;Cross-references: UNIPROT:Q9RI17; UNIPARC:UPI00000DB399; EMBL:AL109962; PIDN:CAB53145.
A;Experimental source: strain A3(2)
C;Genetics:
A;Gene: SCOEDB:SCJ1.27

Query Match 8.7%; Score 123; DB 2; Length 148;
Best Local Similarity 29.7%; Pred. No. 0.014;
Matches 43; Conservative 18; Mismatches 66; Indels 18; Gaps 2;

QY 97 LIRASAAALVCGALGVHFRPERVVGSTAAALASACCPVALVRPHRVPIGRDAAMIV 156
D 2 LVQALSAALAVVGHKRPFGALIGSLATGLAAHSECVVVVGEENRAGPVALGVG 61
QY 157 EADGSSDYG-----LLGAVMAEARLDSFVRVYTCRQSGVGTGDDVRA 201

Db 62 SAAGAAAVRLPGDFRACGGHRRARLDGLERADA---SPRRTEPYACEFGMLAEGR 118

QY 202 SLDRWLARWQRYPDVVRQSAAVHG 226

Db 119 LLAETLAGMCARYPEVVERSTVHG 143

RESULT 15

H70872

probable cypD protease - Mycobacterium tuberculosis (strain H37RV)

C:Species: Mycobacterium tuberculosis

C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004

C:Accession: H70872

R: Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.

; Connor, R.; Davies, R.; Devlin, K.; Feldwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.

; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A: Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A: Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A: Reference number: A70500; MUID:98295987; PMID:9634230

A: Accession: H70872

A: Status: preliminary; nucleic acid sequence not shown; translation not shown

A: Molecule type: DNA

A: Residues: 1-657 <COL>

A: Cross-references: UNIPROT:O53160, UNIPARC:UPI0000128642, GB:AL021184, GB:AL123456, NIT

A: Experimental source: strain H37RV

C: Genetics:

A: Gene: cypD

C: Superfamily: Enterococcus copper-transporting ATPase cypB; ATPase nucleotide-binding d

F: 25-394/Domain: ATPase transduction domain homology <ATP>

F: 465-608/Domain: ATPase nucleotide-binding domain homology <ATN>

Query Match

Best Local Similarity 8.2%; Score 116; DB 2; Length 657;

Matches 69; Conservative 32; Mismatches 101; Indels 66; Gaps 13;

QY 27 AVDEAA--SRDIPRLIYALPDDPGYAAHGAARKLAAMNAVYAF--TAVEAADRP 81

Db 345 ALDKTGTLTRGIP--RLASVAFLDPNVVDARRLLQLAAAEQSEHPLGRATVAEARRRG 402

QY 82 VKVE-----VEITQER-----PVTSIRASAAAL-----V 107

Db 403 IAIIPPAKDFRAVPCGVAHVNDVEIASPQSYRGAPLAEPLASGATPAIVLLDGV 462

QY 108 CVGAIGV--HHFRPERVGTAAALALSAQCPVAIVPRVPIGRDAWIVVEADGSSDIGV 166

Db 463 AIGVIGLTDLQRPDAVESVAAAMALTAAPVLLTGDN-----GR-AAWRVARNNGITDVRA 517

QY 167 LIGAVMABRLRDSPVRYVTCRQSG-----VQPTGDDVRASLDRLARWQRYPDVRYQ 220

Db 518 AL-----LPEQKVEVVNRLLQAGHQVLLVGDGVNDAPAMAAARAAMVAMGAGADLTLLQ 569

QY 221 SA---AVHGEL--LDVLAGIGRSVHMV 243

Db 570 TADGVTTRDELHTIPTIIGLARQARRV 597

Search completed: March 23, 2006, 05:26:27

Job time: 12.1001 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model
Run on: March 23, 2006, 04:56:43 ; Search time 71.3577 Seconds
(without alignments)
2758.529 Million cell updates/sec

Title: US-10-617-038-21
Perfect score: 1410
Sequence: 1 MNOSHKEPSTIVGIDGSKPA.....GNAYLQEAAGCTLLVVGQOYL 279

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues
Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_05.80:*
1: uniprot_sprotc:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1410	100.0	279	Q7Z244_MYCBO	Q7Z244 mycobacteri
2	1410	100.0	279	O53474_MYCTU	O53474 mycobacteri
3	431.5	30.6	268	Q7TX68_MYCBO	Q7TX68 mycobacteri
4	431.5	30.6	268	P95192_MYCTU	P95192 mycobacteri
5	416	29.5	272	Q7TY74_MYCBO	Q7TY74 mycobacteri
6	416	29.5	272	O06188_MYCTU	O06188 mycobacteri
7	374.5	26.6	252	Q73V37_MYCPA	Q73V37 mycobacteri
8	257	18.2	295	Y2005_MYCTU	Y2005 mycobacteri
9	257	18.2	295	Y2028_MYCBO	Y2028 mycobacteri
10	238	16.9	281	Q7WZ45_9ACTO	Q7WZ45 nonomuraea
11	227	16.1	294	Q7J249_MYCPA	Q7J249 mycobacteri
12	224	15.9	294	Q9R128_STRCO	Q9R128 streptomyc
13	214.5	15.2	294	Q7J262_MYCPA	Q7J262 mycobacteri
14	214	15.2	297	Q7Y75_MYCBO	Q7Y75 mycobacteri
15	214	15.2	297	O06189_MYCTU	O06189 mycobacteri
16	213	15.1	343	Q4NL54_9M1CC	Q4NL54 arthrobacte
17	211.5	15.0	294	Q7Z245_MYCBO	Q7Z245 mycobacteri
18	211.5	15.0	294	O53472_MYCTU	O53472 mycobacteri
19	211	15.0	328	Q9R1Y4_STRCO	Q9R1Y4 streptomyc
20	208	14.8	295	O82NK3_STRCO	O82NK3 streptomyc
21	207	14.7	283	Q9R1Y5_STRCO	Q9R1Y5 streptomyc
22	207	14.7	296	O5YV27_NOCFA	O5YV27 nocardia fa
23	207	14.7	301	Q8NLJ9_CORGL	Q8NLJ9 streptomyc
24	206.5	14.6	290	Q82NM7_STRCO	Q82NM7 streptomyc
25	206	14.6	345	O5YV44_NOCFA	O5YV44 nocardia fa
26	203.5	14.4	293	Q7J261_MYCPA	Q7J261 mycobacteri
27	203.5	14.4	296	O5YTS4_NOCFA	O5YTS4 nocardia fa
28	202	14.3	318	O8F1R8_CORER	O8F1R8 streptomyc
29	196	13.9	293	O5YV77_NOCFA	O5YV77 nocardia fa
30	190	13.5	284	O6A8Y6_PROAC	O6A8Y6 propionibac
31	188	13.3	303	Q4YJ55_CORJK	Q4YJ55 corynebacte

32	183.5	13.0	300	Q9K4H8_STRCO	Q9K4H8 streptomyc
33	182.5	12.9	317	Y1196_MYCTU	Y1196 mycobacteri
34	182.5	12.9	317	Y2019_MYCBO	Y2019 mycobacteri
35	180.5	12.8	301	Q9R146_STRCO	Q9R146 streptomyc
36	180	12.8	300	O6NEJ7_CORDI	O6NEJ7 corynebacte
37	171	12.1	312	Q9K4L5_STRCO	Q9K4L5 streptomyc
38	170.5	12.1	288	Q9R148_STRCO	Q9R148 streptomyc
39	167.5	11.9	144	Q4NDF1_9M1CC	Q4NDF1 arthrobacte
40	157	11.1	321	O8G872_BIFLO	O8G872 bifidobacte
41	154	10.9	345	O8G3U1_BIFLO	O8G3U1 bifidobacte
42	148	10.5	151	Q4NGW3_9M1CC	Q4NGW3 arthrobacte
43	146.5	10.4	284	Q4NL45_9M1CC	Q4NL45 arthrobacte
44	139.5	9.9	289	Q4NE71_9M1CC	Q4NE71 arthrobacte
45	138	9.8	159	Q4NDE9_9M1CC	Q4NDE9 arthrobacte

ALIGNMENTS

RESULT 1	Q7Z244_MYCBO	PRELIMINARY:	PRT:	279 AA.
ID	Q7Z244_MYCBO			
AC	Q7Z244			
DT	01-OCT-2003 (Tremblrel, 25, Created)			
DT	01-OCT-2003 (Tremblrel, 25, Last sequence update)			
DT	01-MAR-2004 (Tremblrel, 26, Last annotation update)			
DE	Hypothetical protein MB2053c.			
GN	OrderedLocustNames=MB2053c;			
OS	Mycobacterium bovis.			
OC	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;			
OC	Corynebacteriineae; Mycobacteriaceae; Mycobacterium;			
CC	Mycobacterium tuberculosis complex.			
NCBI	TaxID=1765;			
ON	[1]			
RP	NCBIOTIDB SEQUENCE.			
RC	STRAIN=AF2122/97;			
RX	MEDLINE=22709107; PubMed=12788972; DOI=10.1073/pnas.1130426100;			
RA	Garnier T., Sigmeleer K., Camus J.-C., Medina N., Mansoor H.,			
RA	Pryor M., Duthoy S., Grondin S., Lacroix C., Monsenpe C., Simon S.,			
RA	Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,			
RA	Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewison R.G.,			
RT	"The complete genome sequence of Mycobacterium bovis."			
RL	Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882 (2003).			
DR	EMBL; BX248341; CAD96906.1; -; Genomic DNA.			
DR	GO; GO:0006950; P:response to stress; IEA.			
DR	InterPro; IPR006015; Usp.			
DR	InterPro; IPR006016; UspA.			
DR	Pfam; PF00582; Usp; 1.			
DR	PRINTS; PR01438; UNWRSLSTRESS.			
KW	Complete proteome; Hypothetical protein.			
SQ	SEQUENCE 279 AA; 29441 MW; 52F240F9FA66AC9 CRC64;			
Query Match	100.0%;	Score 1410;	DB 2;	Length 279;
Best Local Similarity	100.0%;	Pred. No. 7.5e-102;		
Matches 279;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1 MNOSHKEPSTIVGIDGSKPAVQAALMAVDENASRDIPRLLYATEPDGPYAAHGAARK 60			
DB	1 MNOSHKEPSTIVGIDGSKPAVQAALMAVDENASRDIPRLLYATEPDGPYAAHGAARK 60			
QY	61 LAAENAVRAFTAFAVEADRPVKEVEITQERPVTSILIRASAAALVCVAGIGHRRPE 120			
DB	61 LAAENAVRAFTAFAVEADRPVKEVEITQERPVTSILIRASAAALVCVAGIGHRRPE 120			
QY	121 RVGSTAALALSAQCPAIVRPHRPVIGRANATIVENADSSDITGVLLGVMAEARLDS 180			
DB	121 RVGSTAALALSAQCPAIVRPHRPVIGRANATIVENADSSDITGVLLGVMAEARLDS 180			
QY	181 PVRVVTGROSGVGTGPDVVASLDRWLARWQPRPDVRSAAVHGELIDYLAGISGVH 240			
DB	181 PVRVVTGROSGVGTGPDVVASLDRWLARWQPRPDVRSAAVHGELIDYLAGISGVH 240			
QY	241 MVLASDQERHVEQLVGAPGNAYLQEAAGCTLLVVGQOYL 279			

```
Db 241 MVLSTASDQEHVEQLVGAPGNAVLQEAAGCTLLVVGQOYL 279
|||||
RESULT 2
ID 053474 MYCTU PRELIMINARY; PRT; 279 AA.
AC 053474; Q7D7L5;
DT 01-JUN-1998 (TReMBLrel. 06, Created)
DT 01-JUN-1998 (TReMBLrel. 06, Last sequence update)
DT 13-SEP-2005 (TReMBLrel. 31, Last annotation update)
DE Hypothetical protein (universal stress protein family).
GN OrderedLocustNames=MT2087, Rv2028c;
OS Mycobacterium tuberculosis;
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium tuberculosis complex.
OX NCBI_TaxID=1773;
RN [1];
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=H37Rv; PubMed=9634230; DOI=10.1038/31159;
MEDLINE=98295987; PubMed=12218035;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C.M.,
RA Harris D.E., Gordon S.V., Eigmeier K., Gae S., Barry C.B. III,
RA Tekaita F., Badcock K., Bauman D., Brown D., Chillingworth T.,
RA Connor R., Davies R.W., Devlin K., Felwell T., Gentles S., Hamlin N.,
RA Holtroyd S., Hornsby T., Jagels K., Krogh A., McLean J., Moule S.,
RA Murphy L.D., Oliver S., Osborne K., Quail M.A., Rajandream M.A.,
RA Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sultson J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."
RL Nature 393:537-544(1998).
RN [2];
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CDC 1551 / Oshkosh; PubMed=12206494; PubMed=12218035;
MEDLINE=22206494; PubMed=12218035;
RA DOI=10.1128/JB.184.19.5479-5490.2002;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J.D., DeBoy R.T., Dodson R.J., Gwinn M.L., Haft D.H.,
RA Hickey E.K., Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D.,
RA Salzberg S.L., Delcher A.L., Utermack T.R., Weidman J.F., Khouri H.M.,
RA Gill J., Mikha A., Bishai W., Jacobs W.R. Jr., Venter J.C.,
RA Fraser C.M.;
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains."
RL J. Bacteriol. 184:5479-5490(2002).
RA EMBL: BX842578; CAI17242.1; -; Genomic DNA.
RA EMBL: AE000516; AAK46366.1; -; Genomic DNA.
RA EMBL: C70942; C70942.
RA TIGR: MT2087; -.
DR Tuberculins; Rv2028c; -.
DR GO: GO:0006950; P: response to stress; IEA.
DR InterPro: IPR006015; Usp.
DR InterPro: IPR005016; UspA.
DR Pfam: PF00582; Usp; 1.
DR PRINTS: PRO1438; UNVRSLSTRESS.
KM Complete proteome; Hypothetical protein.
SQ SEQUENCE 279 AA; 29441 MW; 52F240F9FE466AC9 CRC64;

Query Match 100.0%; Score 1410; DB 2; Length 279;
Best Local Similarity 100.0%; Pred. No. 7, 5e-102;
Matches 279; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Db 241 MVLSTASDQEHVEQLVGAPGNAVLQEAAGCTLLVVGQOYL 279
|||||
RESULT 3
ID 07TX68 MYCBO PRELIMINARY; PRT; 268 AA.
AC 07TX68;
DT 01-OCT-2003 (TReMBLrel. 25, Created)
DT 01-OCT-2003 (TReMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Hypothetical protein MD3158c.
GN OrderedLocustNames=MD3158c;
OS Mycobacterium bovis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium tuberculosis complex.
OX NCBI_TaxID=1765;
RN [1];
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AF2122/97; PubMed=12788972; DOI=10.1073/pnas.1130426100;
MEDLINE=22709107; PubMed=12788972; PubMed=1130426100;
RA Garnier T., Eigmeier K., Camus J.-C., Medina N., Mansoor H.,
RA Pryor M., Dutfoy S., Grondin S., Lacroix C., Monsenpe C., Simon S.,
RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RT "The complete genome sequence of Mycobacterium bovis."
RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
RA EMBL: BX248345; CAD95250.1; -; Genomic DNA.
RA EMBL: BX248345; CAD95250.1; -; Genomic DNA.
DR GO: GO:0006950; P: response to stress; IEA.
DR InterPro: IPR006015; Usp.
DR InterPro: IPR006015; Usp.
DR InterPro: IPR006015; UspA.
DR Pfam: PF00582; Usp; 1.
DR PRINTS: PRO1438; UNVRSLSTRESS.
KM Complete proteome; Hypothetical protein.
SQ SEQUENCE 268 AA; 28008 MW; EE2CB4F9C9D602AE CRC64;

Query Match 30.6%; Score 431.5; DB 2; Length 268;
Best Local Similarity 40.4%; Pred. No. 1, 6e-25;
Matches 111; Conservative 34; Mismatches 105; Indels 25; Gaps 8;
```


AC P95192; Q7D624;
 DT 01-MAY-1997 (TrEMBLrel. 03, Created)
 DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
 DE Hypothetical protein (Universal stress protein family).
 GN OrderedLocustNames=MT3220, RV3134C;
 OS Mycobacterium tuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
 OC Mycobacterium tuberculosis complex.
 NC NCB1_TaxID=1773;
 RX NCBI [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=H37Rv;
 RX MEDLINE=98295987; PubMed=9634230; DOI=10.1038/31159;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C.M.,
 Harris D.E., Gordon S.V., Eigmeier K., Gae S., Barry C.E. III,
 Tekaia F., Badcock K., Basham D., Brown D., Chillingworth T.,
 Connor R., Davies R.M., Devlin K., Feltwell T., Gentles S., Hamlin N.,
 Holtroyd S., Hornby T., Jagsels K., Krogh A., McLean J., Moule S.,
 Murphy L.D., Oliver S., Osborne J., Quail M.A., Rajandream M.A.,
 Rogers J., Rutter S., Seeger K., Skellon S., Squares S., Squares R.,
 Sulston J.E., Taylor K., Whitehead S., Barrett B.G.;
 RT "Deciphering the biology of Mycobacterium tuberculosis from the
 complete genome sequence."
 RL Nature 393:537-544(1998).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=CDC 1551 / Oshkosh;
 RX MEDLINE=22206494; PubMed=12218036;
 RX DOI=10.1128/JB.184.19.5479-5490.2002;
 RA Pleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 Peterson J.D., DeBoy R.J., Dodson R.C., Gwinn L.A., Haft D.H.,
 Hickey E.K., Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D.,
 Salzberg S.L., Delcher A., Utterback T.R., Weisman J.F., Kouri H.M.,
 Gill J.J., Mikula A., Bishai W., Jacobs W.R. Jr., Venter J.C.,
 Fraser C.M.;
 RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
 laboratory strains."
 RL J. Bacteriol. 184:5479-5490(2002).
 DR EMBL; BX842582; CAB06280.1; -; Genomic_DNA.
 DR PIR; G70645; G70645.
 DR TIGR; MT3220; -.
 DR Tuberculast; RV3134C; -.
 DR GO; GO:0006950; P:response to stress; IEA.
 DR InterPro; IPR006015; Unp.
 DR InterPro; IPR006016; UnpA.
 DR Pfam; PF00582; Unp; 1.
 DR PRINTS; PR01438; UNRS1STRESS.
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 268 AA; 28008 MW; EE2CB4F9C9D602AE CRC64;
 Query Match 30.6%; Score 431.5; DB 2; Length 268;
 Best Local Similarity 40.4%; Pred. No. 1.6e-25;
 Matches 111; Conservative 34; Mismatches 105; Indels 25; Gaps 8;
 QY 1 MNQSHKPSIVYIDGSKPAVOALMAVDEAASRDIPRLYLAIPEDDPGYAAHGAARK 60
 DB 1 MSDDRPAVAVVVIGIDGSRPAATTAALMAVDEAVNRDIPRLYLVIDPQLSAAGCGGCG- 58
 QY 61 LAAENAVRYAFTVAADRPVKVEITQERPVTSILRASAAALVCGAIGVHPRPE 120
 DB 59 -SAAAPALHDSARKVEATGCPVKIETEVLCGRPLTKMQESRSAMLCVSGVGLDHVGR 117
 QY 121 RVGTAALALASQCPVAIVPRH-RVPIGRDAWIVVDEADGSSDIGNLLGAVMAEARLR 178
 DB 118 R-GSVAAATLQASALCPVAIVHSPAPRATTSQVSAVVAEVDN---GVLLHAPFEARLR 172
 QY 179 DSPRAVVTGROSGVGDIGDDV-----RASIGRLMAWQRRYPVPRVQSAVHGGELDY 231
 DB 173 GVPLRAVAVH---AAETPDVDEQSGRLAHVHLSRLLAHMTLYPEVRVDRAIAGGSACRH 229

QY 232 LAGIGRSVHMYVTSASDQEHVEQLVGA--PQNAV 264
 DB 230 LAANAKPGOLFV---ADSHSAHELICGAYOPGCAVL 261
 RESULT 5
 ID Q7Y74 MYCBO PRELIMINARY; PRT; 272 AA.
 AC Q7Y74;
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Hypothetical protein MD2657C.
 GN OrderedLocustNames=MD2657C;
 OS Mycobacterium bovis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
 OC Mycobacterium tuberculosis complex.
 NC NCB1_TaxID=1765;
 RX NCBI [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=AF2122/97;
 RX MEDLINE=22709107; PubMed=12788972; DOI=10.1073/pnas.1130426100;
 RA Garnier T., Eigmeier K., Camus J.-C., Medina N., Mansoor H.,
 Pryor M., Duthey S., Grondin S., Lacroix C., Monempe C., Simon S.,
 Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
 RA Parkhill J., Barrett B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
 RT "The complete genome sequence of Mycobacterium bovis."
 RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
 DR EMBL; BX248343; CAD94842.1; -; Genomic_DNA.
 DR GO; GO:0006950; P:response to stress; IEA.
 DR InterPro; IPR006015; Unp.
 DR InterPro; IPR006016; UnpA.
 DR Pfam; PF00582; Unp; 2.
 DR PRINTS; PR01438; UNRS1STRESS.
 KW Complete proteome.
 SQ SEQUENCE 272 AA; 29400 MW; 3EE3A17B3C5720DB CRC64;
 Query Match 29.5%; Score 416; DB 2; Length 272;
 Best Local Similarity 36.8%; Pred. No. 2.6e-24;
 Matches 98; Conservative 46; Mismatches 116; Indels 26; Gaps 2;
 QY 9 SIVGIDGSKPAVOALMAVDEAASRDIPRLYLAIPEDDPGYAAHGAARKLAENAV 68
 DB 11 TIIVIGIDGSHNAITPAALMGVDEAISRAVPLRLVSVIKPTSPDDY---RDLHAERSL 67
 QY 69 RYATVAEADRPVKVEITQERPVTSILRASAAALVCGAIGVHPRERVGSTAA 128
 DB 68 REAQSAAVEAAGKVKIETDIPRGPAGVPLVSAASDAEMICVSGVIGRYASSTIGSTATE 127
 QY 129 LALSQCPVAIVPRHVRPIGRDAWIVVDEADGSSDIGNLLGAVMAEARLDSPVVTCR 188
 DB 128 LAERHACPVAVMWSKVDOPASDINWIVMTDADNEAVLEIYARERKLDQAPLALGGR 187
 QY 189 QSGVGDTGDDVVRASLDRLMARQRRYPDPVQSAVHGGELDYLAGIGRSVHMYVTSASD 248
 DB 188 PEHLRELRD---GSEFERVQDMMHNRHVDVAVYPTTHTGRLADLHDERVQLAVIGGE 244
 QY 249 QEHVEQLVGAAGNAVLOEAGCTLLIV 274
 DB 245 AGQLARLVGPSGHPVFRHACSVLVV 270
 RESULT 6
 ID O06188 MYCTU PRELIMINARY; PRT; 272 AA.
 AC O06188; Q7D6V6;
 DT 01-JUL-1997 (TrEMBLrel. 04, Created)
 DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
 DE Hypothetical protein (Universal stress protein family).
 GN OrderedLocustNames=MT2629, RV2624C;
 OS Mycobacterium tuberculosis.

```

OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriineae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium tuberculosis complex.
OX NCBI_TaxID=1773;
RN
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=H37Rv;
RX MEDLINE=98295987; PubMed=9634230; DOI=10.1038/31159;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C.M.,
RA Harris D.E., Gordon S.V., Eiglmeier K., Gae S., Barry C.E. III,
RA Tekala F., Badcock K., Basham D., Brown D., Chillingworth T.,
RA Connor R., Davies R.M., Devlin K., Feltham T., Gentles S., Hamlin N.,
RA Holroyd S., Hornsby T., Jagels K., Krogh A., McLean J., Moule S.,
RA Murphy L.D., Oliver S., Osborne K., Quail M.A., Rajandream M.A.,
RA Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sultson J.E., Taylor K., Whitehead S., Barrett B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."
RL Nature 393:537-544(1998).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CDC 1551 / Oshkosh;
RX MEDLINE=22206494; PubMed=12218036;
RX DOI=10.1128/JB.184.19.5479-5490.2002;
RA Fleisichmann R.D., Alland D., Eisen J.A., Carpenter J., White O.,
RA Peterson J.D., DeBoy R.T., Dodson R.J., Gwinn M.L., Haft D.H.,
RA Hickey E.K., Kolonay J.F., Nelson W.C., Umayam L.A., Rasmussen M.D.,
RA Salzberg S.L., Delcher A.L., Uitterlinden T.R., Weidman J.F., Kouri H.M.,
RA Gill J., Mikula A., Bhat N., Jacobs W.R. Jr., Venter J.C.;
RA Fraser C.M.;
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains."
RL J. Bacteriol. 184:5479-5490(2002).
DR EMBL: BX842580; CA080618.1; -; Genomic DNA.
DR EMBL: AB000516; AAK47015.1; -; Genomic DNA.
DR PIR: G70572; G70572.
DR TIGR: MT2699; -.
DR TubercuList; Rv2624c; -.
DR GO: GO:0006950; P:response to stress; IEA.
DR InterPro: IPR006015; Unsp.
DR InterPro: IPR006016; Unsp.
DR Pfam: PF00582; Unsp; 2.
DR PRINTS: PRO1438; UNVRSLSTRESS.
KM Complete proteome.
SQ SEQUENCE 272 AA; 29400 MW; 3EE3A17B3C5720DB CRC64;

Query Match 29.5%; Score 416; DB 2; Length 272;
Best Local Similarity 36.8%; Pred. No. 2,66-24;
Matches 98; Conservative 46; Mismatches 116; Indels 6; Gaps 2;

QY 9 SIIVGIDGSKPAVQALMAVDEAASRDIPRLTYAIEPDDPGVAAHGAARKLAENAV 68
DB 11 TIVIGIDGSAITRALMGDEASRAVPLRVIVITPHTSPDDYD---RDLAHAARSL 67
QY 69 RYATFAVEAADRPVKEVEITQERPTSLIRASAAALVCVGAIGVHFRPERVSTAA 128
DB 68 REAQAASVEAAGKLYKRTIDIPGPAHVLVEASRDAMICVSGVIGRAYSSILGSTATE 127
QY 129 LAASAQCPVAVPRHPRPIGRDAWIVYVADSSDGLVGLGAAVAAKRLDSPRVYTCR 188
DB 128 LAEVAHCPVAVMRSKVDQPAIDIMIVMTDADNDAVEAAREKRLQAPLALGGR 187
QY 189 OSGVGIDGDDVRLASLDKRLARMPQRPYDVQVQAAVHGELLDYLAGRSHMVVLSASD 248
DB 188 PEELRELPD---GEFERAVDDMHRRHPDVAVYPTTTGTGARPLADDERVQLAVIGGGE 244
QY 249 QEHVEQIVGAPGNAVLQEAACITLVV 274
DB 245 AGQLARLVGSGHPVFVHACSVLVV 270

```

```

ID Q73V37_MYCPA PRELIMINARY; PRT; 252 AA.
AC Q73V37;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN OrderedLocustNames=MAP3179c;
OS Mycobacterium paratuberculosis.
OC Bacteria; Actinobacteriidae; Actinomycetales;
OC Corynebacteriineae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium avium complex (MAC).
OX NCBI_TaxID=1770;
RN
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K10;
RA Li L., Bannantine J., Zhang Q., Amornsri A., Alt D., Kapur V.;
RT Submitted (SEP-2003) to the EMBL/GenBank/DBJ database.
DR EMBL: AB017238; AA050727.1; -; Genomic DNA.
DR GO: GO:0006950; P:response to stress; IEA.
DR InterPro: IPR006016; Unsp.
DR Pfam: PF00582; Unsp; 1.
KM Complete proteome; Hypothetical protein.
SQ SEQUENCE 252 AA; 26447 MW; FFDAB63B3BA80B1 CRC64;

Query Match 26.6%; Score 374.5; DB 2; Length 252;
Best Local Similarity 39.4%; Pred. No. 4,1e-21;
Matches 106; Conservative 41; Mismatches 93; Indels 29; Gaps 12;

QY 1 MNSHKEPSTIVGIDGSKPAVQALMAVDEAASRDIPRLTYAIEPDDPGVAAHGAAR 59
DB 1 MSTRVAPVAVLVGSDSRAMHAAVWAIDAVGRDIPRLTYVIDPAGAG--GHGPDTR 58
QY 60 KLAANAENAVYATFAVEAADRPVKEVEITQERPTSLIRASAAALVCVGAIGVHFR 117
DB 59 -LAAALAAALADHRAVDAFAQPVKETEILMGTAARKLEQSSAYMLCTGQGLNHACH 117
QY 118 -RPERVSTAAALALASQCPVAV-PRHPRPIGRDAWIVYVADSSDGLVGLGAAVAA 175
DB 118 GGP-----AITSIVRSALCPVAVVQAPSLPAARVSGVAEVDN-----GTVLRHAFEEA 169
QY 176 RLDDSPRVYTCRQSGVGDGDDVRLASLDKRLARMPQRPYDVQVQAAVHGELLDYLAGL 225
DB 170 RLRG---VGLC---AVGN--PSARVLELRRRLRMRLTPDVQASAVLVGVEQHLRAD 220
QY 236 GRSVHVVLSASDQEHVEQIVGAPGNAVL 264
DB 221 HRAGRLV---TDAYRAALCHA--GHSVL 245

RESULT 8
Y2005 MYCTU STANDARD; PRT; 295 AA.
AC P64921; Q10851.
DT 01-OCT-1996 (Rel. 34, Created)
DT 25-OCT-2004 (Rel. 45, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Hypothetical protein RV2005C/MT2061.
GN OrderedLocustNames=RV2005C; MT2061; ORFNames=MTVCY39.12;
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteriidae; Actinomycetales;
OC Corynebacteriineae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium tuberculosis complex.
OX NCBI_TaxID=1773;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=H37Rv;
RX MEDLINE=98295987; PubMed=9634230; DOI=10.1038/31159;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C.M.,
RA Harris D.E., Gordon S.V., Eiglmeier K., Gae S., Barry C.E. III,
RA Tekala F., Badcock K., Basham D., Brown D., Chillingworth T.,
RA Connor R., Davies R.M., Devlin K., Feltham T., Gentles S., Hamlin N.,
RA Holroyd S., Hornsby T., Jagels K., Krogh A., McLean J., Moule S.,
RA Murphy L.D., Oliver S., Osborne K., Quail M.A., Rajandream M.A.,

```

RESULT 7
Q73V37_MYCPA

RA Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."
RL Nature 393:537-544(1998).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=CDC 1551 / Oshkosh;
RX MEDLINE=22206494; PubMed=12218036;
DOI=10.1128/JB.184.19.5479-5490.2002;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J.D., DeBoy R.T., Dodson R.J., Gwinn M.L., Haft D.H.,
RA Hickey E.K., Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D.,
RA Salzberg S.L., Delcher A., Usterback T.R., Weidman J.F., Khouri H.M.,
RA Gill J., Mikula A., Bishai W., Jacobs W.R. Jr., Venter J.C.,
RA Fraser C.M.;
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL J. Bacteriol. 184:5479-5490(2002).
CC -1- SIMILARITY: To B.subtilis yx1E.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; BX842578; CA98383.1; -; Genomic DNA.
DR EMBL; AE000516; AAK46338.1; -; Genomic DNA.
DR PIR; C70759; C70759.
DR TIGR; MT2061; -.
DR TubercuList; RV2005c; -.
DR InterPro; IPR006015; USP.
DR InterPro; IPR006016; UsPa.
DR Pfam; PF00582; USP; 2.
DR PRINTS; PR01438; UNVRSLSTRESS.
DR Complete proteome; Hypothetical protein.
SQ SEQUENCE 295 AA; 30985 MW; 827D60452E24BE33 CRC64;

Query Match 18.2%; Score 257; DB 1; Length 295;
Best Local Similarity 29.4%; Pred. No. 7.2e-12;
Matches 91; Conservative 43; Mismatches 126; Indels 50; Gaps 9;

QY 1 MNOSHKPPIVIGDGSKPAVQAALMAVDEASRDIPRLYLAIKPD----- 47
DB 1 MSKRKQHGIVGVGDGSLSDAALCMGATDAAMNIPITLVHVNVNADVAITPMPYPETW 60
QY 48 -----DPGYAAGAAARKLAAENAVRYAFTAVEADRPVKEVEITQERPVTSILIRASA 102
DB 61 GVMQEDSG-----RQIVA--NAVKLAKAV--GADRKLKSVKSELVFTVPPTWVEISN 109
QY : 103 AALALVCVGAIGVHFRPRVGVSTAAALSLAQCPSVAIVR-----PHRVP--IGRDAA 152
DB : 110 EAEWVVLGSSGRGALARGLGLSVSSSLVRRAGCPVAIVHSDDAVI PPOHAPVLVIGDGS 169
QY : 153 WI-----VVEADSSDGLVLLGAVMAEARLSDSPVRVVTQKQSGVGDGTDDVRASLDRL 207
DB : 170 PVSELATAVAFDEASRRGVELIAVHAW-----SVEVEVLEPGDPSAVQOEAEISLBERL 224
QY 208 ARWQPRYPDVVQSAAVHGEILLDYLAGRSVHNVVLSASDOEHVE-QLVGAPEGNAVLOE 266
DB 225 AGWQERYPDVPSRVVVCNDRPARKLVQKSASQQLVVVSGHGRGGLTGMLLGSVSNVAVLHA 284
QY 267 AGCTLLVVGQ 276
DB 285 ARVPVIVARQ 294

DT 01-OCT-1996 (Rel. 34, last sequence update)
DT 10-MAY-2005 (Rel. 47, last annotation update)
DE Hypothetical protein Mb2028c.
GN OrderedLocustNames=Mb2028c;
OS Mycobacterium bovis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium tuberculosis complex.
RX NCBI_TaxID=1765;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=AP212/97;
RX MEDLINE=22709107; PubMed=12788972; DOI=10.1073/pnas.1130426100;
RA Garnier T., Sigmeier K., Camus J.-C., Medina N., Mansoor H.,
RA Pryor M., Duthey S., Grondin S., Lacroix C., Monsemp C., Simon S.,
RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RT "The complete genome sequence of Mycobacterium bovis."
RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
CC -1- SIMILARITY: To B.subtilis yx1E.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; BX248341; CAD96881.1; -; Genomic DNA.
DR InterPro; IPR006015; USP.
DR InterPro; IPR006016; UsPa.
DR Pfam; PF00582; USP; 2.
DR PRINTS; PR01438; UNVRSLSTRESS.
DR Complete proteome; Hypothetical protein.
SQ SEQUENCE 295 AA; 30985 MW; 827D60452E24BE33 CRC64;

Query Match 18.2%; Score 257; DB 1; Length 295;
Best Local Similarity 29.4%; Pred. No. 7.2e-12;
Matches 91; Conservative 43; Mismatches 126; Indels 50; Gaps 9;

QY 1 MNOSHKPPIVIGDGSKPAVQAALMAVDEASRDIPRLYLAIKPD----- 47
DB 1 MSKRKQHGIVGVGDGSLSDAALCMGATDAAMNIPITLVHVNVNADVAITPMPYPETW 60
QY 48 -----DPGYAAGAAARKLAAENAVRYAFTAVEADRPVKEVEITQERPVTSILIRASA 102
DB 61 GVMQEDSG-----RQIVA--NAVKLAKAV--GADRKLKSVKSELVFTVPPTWVEISN 109
QY : 103 AALALVCVGAIGVHFRPRVGVSTAAALSLAQCPSVAIVR-----PHRVP--IGRDAA 152
DB : 110 EAEWVVLGSSGRGALARGLGLSVSSSLVRRAGCPVAIVHSDDAVI PPOHAPVLVIGDGS 169
QY : 153 WI-----VVEADSSDGLVLLGAVMAEARLSDSPVRVVTQKQSGVGDGTDDVRASLDRL 207
DB : 170 PVSELATAVAFDEASRRGVELIAVHAW-----SVEVEVLEPGDPSAVQOEAEISLBERL 224
QY 208 ARWQPRYPDVVQSAAVHGEILLDYLAGRSVHNVVLSASDOEHVE-QLVGAPEGNAVLOE 266
DB 225 AGWQERYPDVPSRVVVCNDRPARKLVQKSASQQLVVVSGHGRGGLTGMLLGSVSNVAVLHA 284
QY 267 AGCTLLVVGQ 276
DB 285 ARVPVIVARQ 294

RESULT 9
Y2028 MYCBO STANDARD; PRT; 295 AA.
AC P64922; Q10851;
DT 01-OCT-1996 (Rel. 34, Created)

RESULT 10
Q7WZ49_9ACTO PRELIMINARY; PRT; 281 AA.
AC Q7WZ49;
DT 01-OCT-2003 (TremblRel. 25, Created)
DT 01-OCT-2003 (TremblRel. 25, last sequence update)
DT 01-MAR-2004 (TremblRel. 26, last annotation update)
DE Hypothetical protein.
OS Nonomuraea sp. ATCC 39727.

OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Streptomyces; Streptomyces; Streptomyces; Streptomyces; Streptomyces;
 OC NCBI_TaxID=93944;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=ATCC 39277;
 RX MEDLINE=22721466; PubMed=12837387; DOI=10.1016/S1074-5521(03)00120-0;
 RA Soto M., Stinch S., Beltrami F., Lazzarini A., Donadio S.;
 RT "The gene cluster for the biosynthesis of the glycopeptide antibiotic
 A40926 by *Streptomyces* sp.";
 RL Chem. Biol. 10:541-549(2003).
 DR EMBL: AJ561198; CAD91237.1; -- Genomic DNA.
 DR GO: GO:0006950; P:response to stress; IEA.
 DR InterPro: IPR006015; Usp.
 DR Pfam: PF00582; Usp; 2.
 DR PRINTS: PRO1438; UNIVRSSTRESS.
 KW Hypothetical protein.
 SQ SEQUENCE 281 AA; 29593 MW; FGD30F85DF7D41DA CRC64;

Query Match 16.9%; Score 238; DB 2; Length 281;
 Best Local Similarity 28.2%; Pred. No. 2.1e-10;
 Matches 82; Conservative 44; Mismatches 121; Indels 44; Gaps 8;

QY 10 IVVGIDGSKPAVQALMAVDEAASRDIPRLILYAIPEPDDPGYAAHGAARKLAANAENAVR 69
 DB 7 IIVGVDSRTGLAEAGWAAEAQROEPLTVAAHV-----PKWCEG-----AVR 51
 QY 70 YAFPAV-----EAADRPVKEVEITQ-----RPVTSILRASAALAVCGAI 112
 DB 52 YAEVQOMRGAETVLAACDRAKRWPOYSMEVLLPGDRSLVKASEGATILVGSR 111
 QY 113 GVNHFRPERVSTAAALALSNOCFAIVRPHRVEIGRDAWIVVEADGSSDITGLGAVM 172
 DB 112 GIGGRLVGVSVAGVA--AAHPTNVLHDP--AAPRAEVVAGTGSPPCAARAMEFAF 168
 QY 173 AEARLRDSPRVVTC-----RQSGVGTGDVRSLSRMLAKRQPRYDVAVQSAAY 224
 DB 169 AEAELRGRLRAVQAWAMPHPGCFEPASRESEDTLSKESLVHREHGHVAVVAEV 228
 QY 225 HGELLDVLAGRSVHVNL--SASDQEHVEDLVAPGNNAVLOEAGCTLLV 274
 DB 229 HGHPEVLEKARADLVLVGSGHGLAGMLIGISGILLHAPCPMAVV 279

RESULT 11
 ID Q73249_MYCPA PRELIMINARY; PRT; 294 AA.
 AC Q73249;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Hypothetical protein.
 GN OrderedLocustNames=MAP1754C;
 OS Mycobacterium paratuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium;
 OC Mycobacterium avium complex (MAC).
 OC NCBI_TaxID=1770;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=K10;
 RA Li L., Bannantine J., Zhang Q., Amons A., Alt D., Kapur V.;
 RL Submitted (SPP-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL: ARI17333; AAS04071.1; -- Genomic DNA.
 DR GO: GO:0006950; P:response to stress; IEA.
 DR InterPro: IPR006015; Usp.
 DR Pfam: PF00582; Usp; 2.
 DR PRINTS: PRO1438; UNIVRSSTRESS.
 KW Complete proteome.
 SQ SEQUENCE 294 AA; 30978 MW; 0BE0A2A917320CA2 CRC64;

Query Match 16.1%; Score 227; DB 2; Length 294;
 Best Local Similarity 26.7%; Pred. No. 1.6e-09;
 Matches 79; Conservative 44; Mismatches 125; Indels 48; Gaps 9;
 QY 10 IVVGIDGSKPAVQALMAVDEAASRDIPRLILYAIPEPDDPGYAAHGAARKLAANAENAVR 65
 DB 10 IAAVADGSPASNAAYAAAEAAHRCPLITVIANVPTTMYPPVPEALATNLEDECK 69
 QY 66 NAVRYAFVAVE---AADRPVKEVEITQERPVTSILRASAALAVCGAIGVHFRPERV 122
 DB 70 QALHATKIAEAMPADRPQPIGRKILYSAVPALLETSGAVMVLVSGHGLARGL 129
 QY 123 GSTRPAALASOCPVAVR-----PHRVPI-----GRDAWIVVER--DGSSDITGLG 169
 DB 130 GSVSSAVVRHANCFAVAVRDEELDPHSAPVLLCTDGSPPSELTATLAFDEARRGVDLV 189
 QY 170 AVMA-----EARLRDSPRVVTCRQSGVGTGDVRSLSRMLAKRQPRYDVAVR 218
 DB 190 AIAHMSPTAVTEVPEIDWPI-----VEAEERHLAESLAGMERYPDVT 233
 QY 219 VQSAAVHGEILDVLAGRSVHVNL--SASDQEHVEDLVAPGNNAVLOEAGCTLLV 273
 DB 234 VHLVLRADRAAQHIISSSTRAQLVVGSHGRGLARLLGSGVNAVLSRVVPVIV 289

RESULT 12
 ID Q9R128_STRCO PRELIMINARY; PRT; 294 AA.
 AC Q9R128;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Hypothetical protein SC00167.
 GN OrderedLocustNames=SC00167; ORFNames=SCJ1.16c;
 OS Streptomyces coelicolor.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Streptomyces; Streptomyces; Streptomyces; Streptomyces;
 OC NCBI_TaxID=1902;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=A3(2) / M145;
 RX MEDLINE=21996410; PubMed=12000953; DOI=10.1038/417141a;
 RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kleiser H.,
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
 RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
 RA Huang C.-H., Kleiser T., Larke L., Murphy L.D., Oliver K., O'Neill S.,
 RA Rabinowitsch E., Rajandream M.A., Rutherford K.M., Rutter S.,
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
 RA Warren T., Wietzorrek A., Woodward J.R., Bartell B.G., Parkhill J.,
 RA Hopwood D.A.;
 RT "Complete genome sequence of the model actinomycete *Streptomyces*
 coelicolor A3(2).";
 RL Nature 417:141-147(2002).
 DR EMBL: ALJ39104; CAB53134.1; -- Genomic DNA.
 DR PIR: T36949; T36949.
 DR GO: GO:0006950; P:response to stress; IEA.
 DR InterPro: IPR006015; Usp.
 DR Pfam: PF00582; Usp; 2.
 DR PRINTS: PRO1438; UNIVRSSTRESS.
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 294 AA; 31013 MW; 2821576B9B8EA7F CRC64;

Query Match 15.9%; Score 224; DB 2; Length 294;
 Best Local Similarity 26.7%; Pred. No. 2.7e-09;
 Matches 76; Conservative 44; Mismatches 147; Indels 18; Gaps 5;
 QY 10 IVVGIDGSKPAVQALMAVDEAASRDIPRLILYAI--EPDDPGYAAHGAARKLAANAEN 66
 DB 8 LVVGVDGSDGSLAIDMAVDEAQRQGLPRLVYASLWERYEGALLPMGRGRSPBQVAAEN 67
 QY 67 AVRYAFPAVADRPVKEVEITQERPVTSILRASAALAVCGAIGVHFRPERVGSTA 126

Db 68 IVGTAERVRRRYDPLGLTDTDTVPAEAVSALLAEGRHATAVVTSGRGGELKAGLLGSVS 127
Qy 127 AALALSAQCPVAIVRPHRPVIGRDAAMIIVEADSSDIGVLLGAMAEARLRDSPVRVVT 186
Db 128 LAVASRADCPVVVVRGKESKALSISHERVLLGAGPDTTSGAAVRFRRADVRGCELDVVR 187
Qy 187 C-----RQSGVGDGTDD-----VRASLDRLARWQPRYPDVVRQSAAVHGGELLDYL 232
Db 188 AMRCPAYENADEGAPSDSEDPERRBASALIDTLVAEAAAEHPSVRLAKTTIEGPARKVL 247
Qy 233 AGLRSVHMVVLASDQE-HYEQLVGARGNAVLEAGETLLVVGQ 276
Db 248 VHRTAADLVVVGARHRSGRFGLQGRVTHTLQHAACPVAIVVQ 292

RESULT 13

Q73262_MYCPA PRELIMINARY; PRT; 294 AA.
ID Q73262_MYCPA
AC Q73262;
DT 05-JUL-2004 (Tremblrel. 27, Created)
DT 05-JUL-2004 (Tremblrel. 27, Last sequence update)
DT 05-JUL-2004 (Tremblrel. 27, Last annotation update)
DE Hypothetical protein.
GN OrderedLocustNames=MBP1741c;
OS Mycobacterium paratuberculosis.
OC Bacteria; Actinobacteria; Actinomycetales; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium avium complex (MAC).
OX NCBI_TaxID=1770;
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=K10;
RA Li L., Bannantine J., Zhang Q., Amonsin A., Alt D., Kapur V.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE017233; AA04058.1; -; Genomic DNA.
DR GO; GO:0006950; P:response to stress; IEA.
DR InterPro; IPR006015; Usp.
DR Pfam; PF00582; Usp; 2.
DR PRINTS; PRO1438; UNVS1STRESS.
KW Complete proteome.
SQ SEQUENCE 294 AA; 31246 MW; B86F65074298A658 CRC64;

Query Match 15.2%; Score 214.5; DB 2; Length 294;
Best Local Similarity 28.0%; Pred. No. 1.5e-08;
Matches 86; Conservative 47; Mismatches 123; Indels 51; Gaps 12;

Qy 1 MNQSHKPSIVIGIDGSKPAVOALMAVDEASRDIPRLLYAIE-----PDDEGYAAG 55
Db 1 MAASKKGCVLVGVDSPPASNFACMARDAENVPPLTVHMVNAATVMPQVE-MAAEA 59
Qy 56 AAARKLAANAENAVRYAFTAVEAD-----RPVKVEVETIOERPYTSLIRASAAALVCV 109
Db 60 VAMQE-----DGRGVLVQEAVKIADATNGRKLITTELMHAPAPPTLAQISEEAELVVV 115
Qy 110 GAIGVHFRPRVSGSTAALALSACCPVAIVR-----PHRVPI-----GRDAWI-- 154
Db 116 GSTRGAIGRILLSSVSSGLVRRAKCPVAIVHDEDPMPYFORAPVLGVIGIDGSPASIELAT 175
Qy 155 VVEADGSSDIGVLLGAVMAEARLRDSPVRVVTCRQSGVGDGTDD-----VRASLDRLAR-- 209
Db 176 AIAFDEASRRGVDLNAVHAMS---DTQV-----FGLGIDIMPVRSSEARSLAERL 223
Qy 210 --WQPRYPDVVRQSAAVHGGELLDYLAGRSVHMVVLASDQEHVE-QLVAGPAGNVLQOE 266
Db 224 AGMQRERPDVTVHMMVVCDRPARQLIEQSSAQLTVVSGHSGGLAGTLLGSVENAVVHS 283
Qy 267 AGCTLLV 273
Db 284 VRMPVIV 290

RESULT 14
Q77Y75_MYCBO PRELIMINARY; PRT; 297 AA.
ID Q77Y75_MYCBO
AC Q77Y75;
DT 01-OCT-2003 (Tremblrel. 25, Created)
DT 01-OCT-2003 (Tremblrel. 25, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Hypothetical protein TB31.7.
GN Name=TB31.7; OrderedLocustNames=MB2656;
OS Mycobacterium bovis.
OC Bacteria; Actinobacteria; Actinomycetales; Mycobacteriaceae; Mycobacterium;
OC Corynebacterineae; Mycobacterium complex.
OX NCBI_TaxID=1765;
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=AF2122/97;
RC MEDLINE=22709107; PubMed=12788972; DOI=10.1073/pnas.1130426100;
RA Garnier T., Elgimeier K., Camus J.-C., Medina N., Mansoor H.,
RA Pryor M., Duthey S., Grondin S., Lacroix C., Monsemp C., Simon S.,
RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA Parkhill J., Barrrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RT "The complete genome sequence of Mycobacterium bovis";
RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882 (2003).
DR EMBL; BX248343; CAD94841.1; -; Genomic DNA.
DR GO; GO:0006950; P:response to stress; IEA.
DR InterPro; IPR006015; Usp.
DR Pfam; PF00582; Usp; 2.
DR PRINTS; PRO1438; UNVS1STRESS.
KW Complete proteome.
SQ SEQUENCE 297 AA; 31652 MW; A590F7058D1E8695 CRC64;

Query Match 15.2%; Score 214; DB 2; Length 297;
Best Local Similarity 27.2%; Pred. No. 1.6e-08;
Matches 84; Conservative 37; Mismatches 134; Indels 54; Gaps 10;

Qy 1 MNQSHKPSIVIGIDGSKPAVOALMAVDEASRDIPRLLYAIEPD-----DPGY 51
Db 1 MSSGNSSLGIIIVIGIDSPAAQVAVRMAARDAELRKITLVHVASPVAVTWLEVPPLPGV 60
Qy 52 AA---HGAARKLAAANAENAVRYAFTAVEAD---RPVKVEVETIOERPYTSLIRASAAA 104
Db 61 LRWQODHG-----RHLIDDLAKVVEQASLRAGPPTVSHSIVPAAAPTLVDMSKDA 111
Qy 105 ALVVGAGVGHFRPRVSGSTAALALSACCPVAIV-----RPHRVVIGRDAATV 156
Db 112 VLMVVGCLSGGRWEGRLLSVSSGLRHAHCPTVITHDEDSVMPHPQAP-----VLV 164
Qy 157 EADGSSDIGVLLGAVMAEARLRDSPVRVVTCTQSGVGDGT-----GDVVRAS-----LDR 205
Db 165 GVDSSASSELATIAIPFASRRN--VDLVALHMSDVTSEWPGIDMPAQSMAGVLAEL 222
Qy 206 WLAQWQPRYPDVVRQSAAVHGGELLDYLAGRSVHMVVLASDQ-EHVEQLVGAPGNVAVL 264
Db 223 RLAGMQRERPVNAITRVVADQPARQLVGRSBEAQLVVVSGRRGAGLGVSGERTVA 282
Qy 265 QEAGCTLLV 273
Db 283 QLARTPVIV 291

RESULT 15

Q06189_MYCTU PRELIMINARY; PRT; 297 AA.
ID Q06189_MYCTU
AC Q06189; Q7D6V7;
DT 01-JUL-1997 (Tremblrel. 04, Created)
DT 01-JUL-1997 (Tremblrel. 04, Last sequence update)
DT 13-SEP-2005 (Tremblrel. 31, Last annotation update)
DE Hypothetical protein TB31.7 (Universal stress protein family).
GN Name=TB31.7; OrderedLocustNames=MT2698, RV2623;
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinomycetales;

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using SW model

Run on: March 23, 2006, 05:24:53 ; Search time 19.0287 Seconds
(without alignments)
1212.194 Million cell updates/sec

Title: US-10-617-038-21
Perfect score: 1410
Sequence: 1 MNOSHKPSIVVIGIDSKFA.....GNAVLOAGCTLLVVGQOYL 279

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:
1: /cgn2_6/ptodata/1/1aa/5_COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/6_COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/H_COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/PCetus_COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/RE_COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfilest.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	207	14.7	301	2	US-09-605-703B-1870
2	115	8.2	3562	2	US-09-679-279-14
3	107	7.6	815	2	US-09-252-991A-27251
4	106.5	7.6	3546	2	US-09-679-279-13
5	105.5	7.5	426	2	US-09-489-039A-9821
6	102	7.2	458	2	US-09-252-991A-28857
7	102	7.2	1650	2	US-09-252-991A-21798
8	100	7.1	1277	2	US-09-602-787A-54
9	100	7.1	2012	2	US-09-602-787A-50
10	99.5	7.1	1927	2	US-10-152-886-63
11	99	7.0	1138	2	US-09-252-991A-20291
12	99	7.0	2162	2	US-09-477-962-97
13	98.5	7.0	381	2	US-09-252-991A-26281
14	98	7.0	323	2	US-09-252-991A-17111
15	97	6.9	348	2	US-09-489-039A-10294
16	97	6.9	346	2	US-09-252-991A-32334
17	97	6.9	736	2	US-09-252-991A-32256
18	96.5	6.8	849	2	US-09-252-991A-21865
19	96.5	6.8	1996	1	US-08-804-227C-9
20	96.5	6.8	1996	1	US-08-804-198-3
21	96.5	6.8	3491	1	US-07-642-734C-2
22	96.5	6.8	3491	2	US-08-439-009A-2
23	96	6.8	433	2	US-09-351-150A-19
24	95.5	6.8	280	2	US-09-328-352-475
25	95.5	6.8	3170	1	US-07-642-734C-5
26	95.5	6.8	3170	2	US-08-439-009A-5
27	95.5	6.8	4472	1	US-08-804-227C-2

28	95	6.7	1721	2	US-10-042-665A-6	Sequence 6, Appl
29	94.5	6.7	655	2	US-09-252-991A-31645	Sequence 31645, A
30	94.5	6.7	659	2	US-09-902-540-16383	Sequence 16383, A
31	94.5	6.7	735	2	US-09-252-991A-30569	Sequence 30569, A
32	94	6.7	367	2	US-09-902-540-15067	Sequence 15067, A
33	94	6.7	613	2	US-09-758-759-151	Sequence 151, App
34	94	6.7	628	2	US-09-252-991A-29296	Sequence 29296, A
35	94	6.7	1271	1	US-08-095-734-2	Sequence 2, Appl1
36	94	6.7	1271	1	US-08-444-623-2	Sequence 2, Appl1
37	94	6.7	1271	2	US-08-471-869-2	Sequence 2, Appl1
38	94	6.7	1271	2	US-09-342-563-2	Sequence 2, Appl1
39	94	6.7	1271	4	PCT-US94-08267-2	Sequence 2, Appl1
40	93.5	6.6	293	2	US-09-252-991A-22215	Sequence 22215, A
41	93.5	6.6	426	2	US-09-902-540-11847	Sequence 11847, A
42	93.5	6.6	569	2	US-09-252-991A-20194	Sequence 20194, A
43	93	6.6	407	2	US-09-252-991A-19349	Sequence 19349, A
44	93	6.6	3567	1	US-07-642-734C-4	Sequence 4, Appl1
45	93	6.6	3567	2	US-08-439-009A-4	Sequence 4, Appl1

ALIGNMENTS

```

RESULT 1
US-09-605-703B-1870
; Sequence 1870, Application US/09605703B
; Patent No. 6962989
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Krogel, Burkhard
; APPLICANT: Schröder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauser, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING NOVEL
; FILE OF INVENTION: PROTEINS
; FILE REFERENCE: BGI-129CB
; CURRENT APPLICATION NUMBER: US/09/605, 703B
; CURRENT FILING DATE: 2000-06-27
; PRIOR APPLICATION NUMBER: 60/142,764
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: 60/152,318
; PRIOR FILING DATE: 1999-09-03
; NUMBER OF SEQ ID NOS: 2934
; SEQ ID NO 1870
; LENGTH: 301
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-605-703B-1870
Query Match 14.7%; Score 207; DB 2; Length 301;
Best Local Similarity 26.1%; Pred. No. 8.2e-14;
Matches 77; Conservative 46; Mismatches 14; Indels 38; Gaps 10;
QY 10 IVGIDSKPAVQALMAVDEASRDIPRLI-----LYA---IEPDDGYAANGAA 57
DB 7 VVAVVDSDSKQAVRWANATANKRGIPRLIASYTPQGFYAGWVPDELFPDLOAEA 66
QY 58 ARKLAAMNVRVYFTVAEADRPVXVEVETTORPPTSILIRAAALVCGAIGVHHF 117
DB 67 LEKINEARD-----IAHEVAPF-IKIGHTAESSPIDMLEMSPDITMIVSGRGJGL 119
QY 118 RPEVSTAAALALSAOCVAVIVRPHRVPIGRDAW--IIVADGSSDIDGVLGAVMAEA 175
DB 120 SGWMSVSGAVSHACFPVVVAREDA-VNEDSKTGPVVGVGDSVSGQATYAFAEA 178
QY 176 RLR-----DSPRVVTCRQSGVGTGDV--BASLDRLIARQP--RYPDVVQ 220
DB 179 EARGAEIVAVHTWMDQVQSLAGIAAQOQWDEVERQOTDMLIERLAPVEKYPSTYVK 238
QY 221 SAHNGELDVLTLAIGRSYHNVVLSASDQEHVE-QLVGACGNVLOAGCTLLIV 274
DB 239 KIITRDRPVRALAEASBNADLLVVGSHRGQFGKMLIGSTSRALLQSAFCPMNVV 293

```

```
RESULT 2
US-09-679-279-14
; Sequence 14, Application US/09679279
; Patent No. 6524841
; GENERAL INFORMATION:
; APPLICANT: McDaniel, Robert
; TITLE OF INVENTION: Recombinant Megalomicin Biosynthetic
; FILE REFERENCE: 30622004700
; CURRENT APPLICATION NUMBER: US/09/679,279
; CURRENT FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/158,305
; PRIOR FILING DATE: 1999-10-08
; PRIOR APPLICATION NUMBER: US 60/190,024
; PRIOR FILING DATE: 2000-03-17
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 3562
; TYPE: PRT
; ORGANISM: Micromonospora megalomicea
US-09-679-279-14
```

```
Query Match      8.2%; Score 115, DB 2, Length 3562;
Best Local Similarity 27.4%; Pred. No. 0.023;
Matches 95; Conservative 25; Mismatches 119; Indels 108; Gaps 20;
```

```
QY 13 GIDSKPAVQAALMAVDEAASRDIPRLRLVA-----IERDDPGYAAH-GA----- 56
D 2743 GDDVPRPGA-AAVGVRCQAQESPDRFLVDGPETPRPAVDPQLVRDGAAPVPRLT 2801
QY 57 --AAKLAAMENAVR-----YAFVAEADRP-----VKVEVEITQER----- 92
D 2802 PLGEPVAVADRAVRLVPGNGSGSIEAFAFVPPADRPPLAEERVAVRAIGVFRDVL 2861
QY 93 -----PYTSLIRASAAAALVCVGAIGVHHFRPERVGSAAALALSAOCPPAIVPRH--- 144
D 2862 ALGMYPRPAEMGTASGVTEVGS-GVRRFTP---GQAVTGLFGAGGPVA-VADHRLLT 2916
QY 145 -VPIGRDAAMTVNA-----DSSSDI-----GVTLGAVM----- 172
D 2917 PVPBG---MAVDAVAAVPIAFTTAHYALHDLAGLQAGSVLVHAAAGVGMAAVLARR 2972
QY 173 --AEARLRDSFVRVVTGROSGVGTGDDVRASL-----DRWLARWQPRYPDVVQSAVH 225
D 2973 AGAEVPAATASPAKPIRLALGLD--DHIASSRESGGERFAARTGKRGVDVVLNS--LT 3028
QY 226 GELID---YIAGIGRSVHVHVLTSASDQEHVEQLVGAPGNAVLOEAG 268
D 3029 GDLDIESARLADGDGVFVEM---GKTDLRPAEQFRGRYVPPDLAEAG 3072
```

```
RESULT 3
US-09-252-991A-27251
; Sequence 27251, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27251
; LENGTH: 815
; TYPE: PRT
```

```
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27251
```

```
Query Match      7.6%; Score 107, DB 2, Length 815;
Best Local Similarity 26.9%; Pred. No. 0.019;
Matches 87; Conservative 31; Mismatches 107; Indels 98; Gaps 19;
```

```
QY 4 SHKPSIVGIDDSKPA-----VQAALMAVDEAASRDIPRLRL-----LVAIERDDP 49
D 293 SLSEPLAAMAAALAGLAPRGGRWRRLAQALLAVLGTGAVVHLRLRORPLDPAARAGP 352
QY 50 GY-----AAGAAARKLALEN-----AVRVAFTVAEADRPVKVEVEITQERPVY 95
D 353 GVARRRRAAAGAAQRLAAAMRWGLPGETDPIAR-AVRPLAAGLARAEV---HRRAG 408
QY 96 SLIRASAAAALVCVGAIGVHHFRPERVGS-----AAALAL--ACPPAIVR----- 141
D 409 QRRRLAABAAAGVLAEEAVLGSVLQRRRAGATGGGQGRRLVASRRPAGDDPLAIVRLARRG 468
QY 142 -----PHRVPY-GRDAWIVVEADSSDITVLGAVMAEARLRDSFVRVVTGROS 150
D 469 APALALPAGGTVPAMPGRAAMLR-----PRVGAAG-TRADARVVAAGVRA-----A 518
QY 191 GVGDGDDVRASLDRWLARWQPRYPDVVQSAAV-HG-----ELDY---LAGLG 236
D 519 GI-----DGAQAGL-----AVGVQAVVGHAGSEDLALVVEVLDAHLFQSLG 562
QY 237 RSV-HMVVLSASDQEHVEQLVGA 258
D 563 DSLGRLLALELVDSHADQVDA 565
```

```
RESULT 4
US-09-679-279-13
; Sequence 13, Application US/09679279
; Patent No. 6524841
; GENERAL INFORMATION:
; APPLICANT: McDaniel, Robert
; TITLE OF INVENTION: Recombinant Megalomicin Biosynthetic
; FILE REFERENCE: 30622004700
; CURRENT APPLICATION NUMBER: US/09/679,279
; CURRENT FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/158,305
; PRIOR FILING DATE: 1999-10-08
; PRIOR APPLICATION NUMBER: US 60/190,024
; PRIOR FILING DATE: 2000-03-17
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 3546
; TYPE: PRT
; ORGANISM: Micromonospora megalomicea
US-09-679-279-13
```

```
Query Match      7.6%; Score 106.5, DB 2, Length 3546;
Best Local Similarity 23.9%; Pred. No. 0.18;
Matches 81; Conservative 33; Mismatches 120; Indels 105; Gaps 18;
```

```
QY 19 PAVQAALMAVDEAASR-----DIPRLRLYAE--PDDPGYAAHGAARKLAA-----AENAV 68
D 1425 PFGREPRFWLEPNPARRVADSDVSLRYRLIEMHTPBGEGRGIDGTWLLTYTGRADDRV 1484
QY 69 RVAFTVAEADRPVKVEVEITQERP-----VTSIRASAAAALVCVGAIGVHHFRPERVG 123
D 1485 EAARQALIESAG--ARVEDLVYERTGRVDLVRLLDAGVPVAGVLCLEPVA-----EPAAEH 1538
QY 124 STAAALALS-----ACPPAIVRPRVPIG-----RD-----AAW-----IVV 156
D 1539 SPLAVTSLSDTLTLTQAVAGSGRECPIMVYTTENAIVAVGPERLRDPAHGALMLGRVAL 1598
QY 157 E-----ADGSSDITGVLL-GAVMAEARLR-DSFVRVVTGROSGVGTGDD 198
```


Db 1599 ENPAWGGGLVDVPGSGVAELSHLGTLSGAGEBDVALRPGDYARMCW-AGAGGTG-- 1655
Qy 199 VRASLDNR-----LARMQPRDYDVQSAAVHG-----ELLDYLA 233
Db 1656 -----KQPRGTGLVYTGCTGCGVGHVAVRMLARQGTPLCLVLNRRGPPADGVEELTTELA 1709
Qy 234 GLGRSVHNVVLASDQEHVEQLVGA-----PGNAVLOEA 267
Db 1710 DLGTRATVTACDVTDREQLRALLATVDEHPLSAVFHVA 1748

RESULT 5
US-09-489-039A-9821

; Sequence 9821, Application US/09489039A
; Patent No. 6610936
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 9821
; LENGTH: 426
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-9821

Query Match 7.5%; Score 105.5; DB 2; Length 426;
Best Local Similarity 23.6%; Pred. No. 0.011; Matches 53; Conservative 38; Mismatches 93; Indels 41; Gaps 8;

Qy 57 AARKLAAMNAVRYAFTAVEADRPVKEVEITQRPVTLIRASAAALVCVGAIGVHH 116
Db 36 AGVSLAAGMGEVEVAFGTGQPS--TARALINSDRMTAAADKVALRGAATFVPH 93
Qy 117 FRPERVGT-----AAALASQCPVALRPHRVITGDAAMIVVADGSSDIGVLL 168
Db 94 WQVKKISTLTGNGPGELEAMMAAGCRMVAVAPAYPAAGHRT-----RDGRCYVH--- 144
Qy 169 GAWAABARLRDSPRVVTCROSGVGTGDDVRASLDRLARMQPRDYDVQSAAVHGEI 228
Db 145 GVPIDQTEFPASDPKTPVS-----PAETSEITIA-MQSKLPCLTINA---GQL 186
Qy 229 LDVLAGLGRSVHNVVLASDQEHVEQLVGAPEGNAVLOEAGCTLLV 273
Db 187 PALATAGEEKRVLLVDMDSHLDQV-----DANAPHARETLIV 227

RESULT 6

US-09-252-991A-28897
; Sequence 28897, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 28897
; LENGTH: 458
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-28897

Query Match 7.2%; Score 102; DB 2; Length 458;
Best Local Similarity 22.6%; Pred. No. 0.029;
Matches 75; Conservative 42; Mismatches 117; Indels 98; Gaps 16;

Qy 27 AVDEASRDIFL--RLIYALEPDDPGYAAHGAARKLAAMENAVRYAFTAVEADRPVK 84
Db 26 AATSSATRAFAALATSLLYACAAPASWTASQSPQVWMEDEFLFENVPALHDQTVRQ 85
Qy 85 EVEIT-----QRPV-----TSLIRASAAALVCVGAIGVHHRRPRV----- 122
Db 86 LARISLGSTRVRYLVANTYGRPRIRGRATLAKPGRSAVAATGSLDVTGGRRTATIA 145
Qy 123 -GST-----AAALASQCPVALRPHRVPI-----GRDAAMIV-----E 157
Db 146 PGATLLDPPVAMPFPAALSOVLVSLYLAAFTMEFTHDGTQGTGIVAGEQTSALQLDA 205
Qy 158 ADG-SSDIGVLGAWMAEARLRDSPRVVTCROSGVGT--GDVRSALDRMLARMQPRYP 215
Db 206 ADSQTSARILLTGILVET--ETATRTLVT---LQDSITDGASASLDR-----NSRWP 253
Qy 216 DVRQSAVHG-----ELL-----DYLGRSVHNVVLASDQE 250
Db 254 DFLAERLAPHGVAVAVNAGISGARLLSDMGASALALRDVLAQPGASSWVVMGLIND-- 311
Qy 251 HVEQLVGAPGNAVLOEAG---CTLLVVGQOYL 279
Db 312 -----IAMPGTAFARNNAAPPLENLINGYROL 338

RESULT 7

US-09-252-991A-21798
; Sequence 21798, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 21798
; LENGTH: 1650
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-21798

Query Match 7.2%; Score 102; DB 2; Length 1650;
Best Local Similarity 23.5%; Pred. No. 0.18;
Matches 87; Conservative 28; Mismatches 114; Indels 142; Gaps 19;

Qy 10 IVVGID-GSKPAVQDALMAVDEASRDIFRL-----LYAIEPDD----- 48
Db 1184 IATLDHGAVALQOAVLAVQARLDPOFALGEQLAATVVOVSADVPDRPQLQAAALV 1243
Qy 49 -----PGVAHGAARKLAAMENAVRYA-FTAVEAA--DRPVKEVEITQE----- 91
Db 1244 QRLARLAFQRAAGHQAAGVAGAAQAEVAAAGVAPLVRDLNAGPHFPEVNEARGVGRAPFA 1303
Qy 92 ----RPVTSIRASAAALVCV-ALGVNHF-REERV-----GSTAALASQCPVALIV- 140
Db 1304 RQGLLPVDQFLRAADLGAATGVGLAVQDLHGGRDLGVALAGLGAQQLAAGVDADVAAG 1363
Qy 141 -----RPHRVIPGRDA-----WIVVADGSS----- 162
Db 1364 GTVAGQPHPAL--LGADQADRPGVHPAQRRAVDRQLMLAALIGARGQLAVGLDVVAAG 1422

```

P07689      1 PRIOR APPLICATION NUMBER: DE 19940765.7
P07689      2 PRIOR FILING DATE: 1999-08-27
P07689      3 PRIOR APPLICATION NUMBER: DE 19940766.5
P07689      4 PRIOR FILING DATE: 1999-08-27
P07689      5 PRIOR APPLICATION NUMBER: DE 19940830.0
P07689      6 PRIOR FILING DATE: 1999-08-27
P07689      7 PRIOR APPLICATION NUMBER: DE 19940831.9
P07689      8 PRIOR FILING DATE: 1999-08-27
P07689      9 PRIOR APPLICATION NUMBER: DE 19940832.7
P07689     10 PRIOR FILING DATE: 1999-08-27
P07689     11 PRIOR APPLICATION NUMBER: DE 19940833.5
P07689     12 PRIOR FILING DATE: 1999-08-27
P07689     13 PRIOR APPLICATION NUMBER: DE 19941378.9
P07689     14 PRIOR FILING DATE: 1999-08-31
P07689     15 PRIOR APPLICATION NUMBER: DE 19941379.7
P07689     16 PRIOR FILING DATE: 1999-08-31
P07689     17 PRIOR APPLICATION NUMBER: DE 19941395.9
P07689     18 PRIOR FILING DATE: 1999-08-31
P07689     19 PRIOR APPLICATION NUMBER: DE 19942077.7
P07689     20 PRIOR FILING DATE: 1999-09-03
P07689     21 PRIOR APPLICATION NUMBER: DE 19942078.5
P07689     22 PRIOR FILING DATE: 1999-09-03
P07689     23 PRIOR APPLICATION NUMBER: DE 19942079.3
P07689     24 PRIOR FILING DATE: 1999-09-03
P07689     25 PRIOR APPLICATION NUMBER: DE 19942088.2
P07689     26 PRIOR FILING DATE: 1999-09-03
P07689     27 NUMBER OF SEQ ID NOS: 678
P07689     28 SEQ ID NO 54
P07689     29 LENGTH: 1277
P07689     30 TYPE: PRF
P07689     31 ORGANISM: Corynebacterium glutamicum
P07689     32 US-09-602-787A-54
P07689     33
P07689     34 Query Match          7.1%; Score 100; DB 2; Length 1277;
P07689     35 Best Local Similarity 21.6%; Pred.No. 0.21;
P07689     36 Matches 55; Conservative 47; Mismatches 113; Indels 40; Gaps 9;
P07689     37
Q07689      1 4 SHKPSYIVGIDGSKPAVQAALMA-----VDEASRDIPLFLVAIEPDDCGYAAHGAA 58
Q07689      2   :|::||::||::||::||::||::||::||::||::||::||::||::||::||::||:
Q07689      3 Db    313 AHTAPDVLVG--RAWPAFAVAVKSAVIPGTSTASAVVEGMSLV-----HLEHHITLV 361
Q07689      4   59 RKLAAENAVYAFTAVEAADRP-----VKVEEITOEPRVTSLI-----RASAAA 105
Q07689      5 Db    362 KSDVPPTDGALKVSATDAEDVDTTDGLRLVIYAETADAEGLNLTIAERFAIRGRGNVA 421
Q07689      6 Qy    106 LVCGAIGVHHFRPERVSGSTAALLALSACFEVAIVRPHRVIG-RDAA-----WIWE 157
Q07689      7 Db    422 RTNTSLAPTPTVDTDRSARAAVATVAPEBMRPFAVISGDRIPIHVSIDVAASLAGLPGVIV 481
Q07689      8 Qy    158 ADGSSDICVL-LGAVMAEARLDSEPVRYTTRQSGVGT--GDVVRALDDMLMARQRY 214
Q07689      9 Db    482 GMWTSAIGELTAGAAFNDEOIQTPAAKVVEYTATMLAFLVGEELIEFSVERSAVDNRDGM 541
Q07689     10 Qy    215 PDVVRVOSAAYHGELL 229
Q07689     11 Db    542 GEVRTVTATVNGNLV 556
Q07689     12
RESULT 9
Q07689     13 / Sequence 50, Application US/09602787A
Q07689     14 / Patent No. 6696561
Q07689     15 / GENERAL INFORMATION:
Q07689     16 / APPLICANT: Pompejun, Mark
Q07689     17 / APPLICANT: Krüger, Burkhard
Q07689     18 / APPLICANT: Schöder, Hartwig
Q07689     19 / APPLICANT: Zeldner, Oskar
Q07689     20 / APPLICANT: Haberhauser, Gregor
Q07689     21 / TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
Q07689     22 / TITLE OF INVENTION: INVOLVED IN MEMBRANE SYNTHESIS AND MEMBRANE
Q07689     23 / TITLE OF INVENTION: TRANSPORT
Q07689     24 / FILE REFERENCE: BGI-125CP
Q07689     25 / CURRENT APPLICATION NUMBER: US/09/602,787A

```

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	524	525
---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

Db 1391 SLHREGVAALTPDOGVAVLRLRLADPEATGVVWSGR 1427

RESULT 11

US-09-252-991A-20291
Sequence 20291, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT FILING DATE: 1999-02-18
PRIORITY FILING DATE: 1999-02-18
PRIORITY FILING DATE: 1998-07-27
PRIORITY FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 20291
LENGTH: 1138
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20291

Query Match 7.0%; Score 99; DB 2; Length 1138;
Best Local Similarity 23.9%; Pred. No. 0.23;
Matches 70; Conservative 26; Mismatches 91; Indels 106; Gaps 15;

QY 44 IEPDDPGY---AAHGAARAKLAAMENAVRYA-----FTAVEADRPVKEVEITQ 91
DB 562 VEPDRGTATPGIRGIAIVLAADVRFEGAORADVQLAAGADGP-----RQ 613
QY 92 RP-----VTSIRASAAALVCV-----AIGVHFRPE-----RVGSTAALAL 131
DB 614 RPSRDRVTTPACIGAGFAADVBPADLANOGALAHVLEEVROGGLDGLADPLAV 673
QY 132 SAQCPVAIVRPHRVPIGRDAMIIVEADSSDIGVLGAVAAEA-----R 176
DB 674 QVLRGLAV--QHR-----GVVGAVALAADHRYEAGEDEVGVQLH 711
QY 177 LRDSRVVVTCTGSGVDGTGDVRA-----LDRML-----ARMOPRPDVAVQSAV 224
DB 712 RRVGVPAVDFAPRAGVAGCGEAGAEHADIFRVLDEVAVGAGGRRRGLGGQCRVDRGEA 771
QY 225 HGEILDVLAGRGSRVHNVLSASQF-----HV-----EQLVG--APGNAV 264
DB 772 GGEVVDHRRHAGRG---ALAAEDRQAAVVVHVGAELGVQGEVAQGEAVL 820

RESULT 12

US-09-477-962-97
Sequence 97, Application US/09477962
Patent No. 6927286
GENERAL INFORMATION:
APPLICANT: SHEN, BEN
APPLICANT: DU, LIANGCHENG
APPLICANT: SANCHEZ, CESAR
APPLICANT: CHEN, MEI
APPLICANT: EDWARDS, DANIEL J.
TITLE OF INVENTION: BLEBOMYCIN GENE CLUSTER COMPONENTS AND THEIR USES
FILE REFERENCE: 407T-895820US
CURRENT APPLICATION NUMBER: US/09/477, 962
CURRENT FILING DATE: 2000-01-05
PRIORITY FILING DATE: 60/115, 435
PRIORITY FILING DATE: 1999-01-06
PRIORITY FILING DATE: 1999-02-05
PRIORITY FILING DATE: 1999-02-05
NUMBER OF SEQ ID NOS: 133
SOFTWARE: PatentIn Ver. 3.0
SEQ ID NO 97
LENGTH: 2162

TYPE: PRT
ORGANISM: Streptomyces verticillius
FEATURE:
OTHER INFORMATION: ORF26
US-09-477-962-97

Query Match 7.0%; Score 99; DB 2; Length 2162;
Best Local Similarity 24.5%; Pred. No. 0.57;
Matches 79; Conservative 35; Mismatches 122; Indels 86; Gaps 16;

QY 7 PPSIVGIDG-----SKPAV-----QAALMAVDEAASDIPRLIYAIEPDDPGYAAH 54
DB 256 PABLVAIDGVAGKSLSTDPATVTHAARIVAKRL--AGERNLPVAL-----TRD 302
QY 55 GAARAKLAAMENAVRYAFTAVEADRPVKEVEITQRPVTSIRASAAALVCVAGIV 114
DB 303 GRSHPELR-----TAIGAERFLPLVHEIRHETAPAEYAR--ALDALVAEGSELL 350
QY 115 HHRPERVGSSTAALALSAQCPVAIVRPHR--VPGR-DAAWIVEADSSDIGVLGAV 171
DB 351 DHCPPELIGS---LDGTAGPCFTFTNHQAEFPVRRAGITFTTVHQDSGTPIPVRLTAR 406
QY 172 MAEARLR-----DSPVRVVTCTGSGV-----GDTGDVRAASLRW--LAR 209
DB 407 RDGARLRMEIGYDEGRIDETFPENAACTRIIEGVVASAPEGVG-DIRMLSDETARLR 465
QY 210 WQRPVDPVAVQSAVHGEILDVLAGRSVHNVLSASDQ-----EIVEQ 254
DB 466 EAGIGRPVELPGNAVHLEFPAEQAA---RTPGNAVAVSAGBDALVYALDEBSNRLAHNLIG 522
QY 255 LVGAPGNALVQERAGCTL-LVVG 275
DB 523 LGVTGPRHVVVSVGRSHELIVG 544

RESULT 13

US-09-252-991A-26281
Sequence 26281, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252, 991A
CURRENT FILING DATE: 1999-02-18
PRIORITY FILING DATE: 1998-07-27
PRIORITY FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 26281
LENGTH: 381
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26281

Query Match 7.0%; Score 98.5; DB 2; Length 381;
Best Local Similarity 26.1%; Pred. No. 0.053;
Matches 82; Conservative 29; Mismatches 122; Indels 81; Gaps 17;

QY 9 SIIVGIDGSKPA--VOAALMAVDEAAS-RDIPRLIYAIEPDDPGYA----- 52
DB 94 ALVVGFGRPARCATVNAHPRDADAGADVPEQLAGAGARQAGHAGADVALGOLAVAI 153
QY 53 -----ANGAARAKLAAMENAVRYAFTAVEADRPVKEVEITQRPVTSIR 100
DB 154 GFVGOARGORAHGAGA--GAAPH--RHQVEIGARQORVAGEAVETALERPDAVFOHA 208
QY 101 SAALAVCVGAIHHRPERVGSSTAALALSA---QCP-----VAIVRPHRVPIGRDAV 153
DB 209 QATGA-----EATGV-----QQAAGDGGRAVAVVAEDQAPAVVULCICORRAAHNORQACD 259

QY 154 IV---VEADSSDIGVLGVAWAEARLRSPKRVVTCRQSGVDTGDDVR-----ASL 203
Db 260 LLOGPAEAGGGGREG-----GGRRODPQFG---RQLPSETGADAEQHRVAAGQHA 307
QY 204 DRWLRLKQPRYPYDVRVQSAAVHGEILDYLAGLGRSVHMYVLSASDQENHQLVGAPGNAY 263
Db 308 GRLLAPGQQWFGGERRRPFAGG-----ADARRQGFOLARRRDPDGLQQ--GTPG--L 357
QY 264 LOEAGCTLLVVGQ 277
Db 358 LAEAGITVLADAHQ 371

```

RESULT 14
US-09-252-991A-1711
: Sequence 1711, Application US/09252991A
: Patent No. 6551795
: GENERAL INFORMATION:
: APPLICANT: Marc J. Rubenfield et al.
: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS.
: TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
: FILE REFERENCE: 107196.136
: CURRENT APPLICATION NUMBER: US/09/252,991A
: CURRENT FILING DATE: 1999-02-18
: PRIOR APPLICATION NUMBER: US 60/074,788
: PRIOR FILING DATE: 1998-02-18
: PRIOR APPLICATION NUMBER: US 60/094,190
: PRIOR FILING DATE: 1998-07-27
: NUMBER OF SEQ ID NOS: 33142
: SEQ ID NO 1711
: LENGTH: 323
: TYPE: PRT
: ORGANISM: Pseudomonas aeruginosa
: US-09-252-991A-1711

```

	Query Match	7.0%	Score 98:	DB 2:	Length 323;
	Best Local Similarity	27.7%;	Pred. No. 0.047;		
	Matches	61; Conservative	22; Mismatches	81; Indels	56; Gaps
Qy	29 DEAASRDIPLRLYLAIIEP-----DDPGYAAHGAARKLA-AAENAVRYAFVAVEADRPV	82			
Db	86 DLAAETDQPRGVFDFGIHLPRVRGVMAGVAHLAAHGQLALVAEHRLAQGLADDAQRSLV	145			
Qy	83 KVEVEITGEREVTLSIRASAAAALVCVALGYH-----HRRPE	120			
Db	146 RTPIQRRQR-----TDAQADFLVVERQVHRQAQRGEERRDRGHQAQVALHY--	197			
Qy	121 RVGSTAAALALS----AACPVALVRPHRVPI-GRDAAMI VVEADSSDIGLLGVMAE-	174			
Db	198 RATEVEAAIALAEPERRARPAIGLVARRHVHGMPEKADATVAGADGGIETG-LAAVGVEE	255			
Qy	175 -----ARLRDSPRVYTCKSGSVC-----DTGDVVAS	202			
Db	. . . 256 QFAVDIAETROVADVDDQRQGLAADVGEHQVGGDVAAA	295			

```

RESULT 15
US-09-489-039A-10294
; Sequence 10294, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Bretton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OR INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489, 039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 10294
;
; LENGTH: 348
; TYPE: PRT

```

```

i ORGANISM: Klebsiella pneumoniae
US-09-489-039A-10294

Query Match      6.9%; Score 97; DB 2; Length 348;
Best Local Similarity 22.7%; Pred. No. 0.067;
Matches 70; Conservative % 42; Mismatches 109; Indels 88; Gaps 17,

14 IDGSKPAVOALMNVDEAASRDIDRLIYAEIRDPGY-----AAHG--AAARKL 61
72 IDGGQVA-----AAAGCDV--VHAANP--PGRRHRCQVLPMLNHTLQAAERKR 118
62 AA-----AENARVAFTAVEADRPV--VVEVEITQERVTSILIRASAAALVCV 109
119 ALVVLPGCTVTVNGNA--FLIAEEAQOQPTRKGAIRVAMELLIEDYVQGGPALLVRA 176
110 GAIGVHHFRPERVSGT--AAALALSAOCPAVIRPHRPVPIGRDAAMITVEADGSSDIGVL 167
177 G-----DFFGRAGNNMFPQGLYKRGQLPRISIPGIIIGVGHQWAL-----PDVLAAT 224
168 LGAVMAERLRIDSFVRVVTCTQSGVGD--TEDDVASLDRLMARQPR-----YDVRVQVS 221
225 IALLLRHHELEPPARF--HMQHMPBDGSEMGQAIQAVVARYGGRVAVVKSFPMWLVKL 281
222 AAV-----HBEILDYLAGLGRSVHNVVLSASDQHNHQQLYGAR 259
282 AAFPNAATIREVVENHLYMRPLVRLRNDKLVDF--LGAEPHTPIDSAYVQY--LQGLGCLP 337
260 GNAVLTQENG 268
338 AGAINTQENG 346

```

Search completed: March 23, 2006, 05:30:09
Job time : 20.1287 secs

THIS PAGE BLANK (USPTO)

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 23, 2006, 06:14:33 ; Search time 58.6719 Seconds
(without alignments)
1986.885 Million cell updates/sec

Title: US-10-617-038-21

Perfect score: 1410
Sequence: 1 MNOSHKPESTIVGIDGSKPA.....GNAVLQEAAGCTLLVVGQOYL 279

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_ML_Main:
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBSCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBSCOMB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBSCOMB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US10_PUBSCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBSCOMB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBSCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1410	100.0	279	4	US-10-617-038-21
2	431.5	30.6	268	4	US-10-617-038-42
3	416	29.5	272	4	US-10-617-038-27
4	214	15.2	297	4	US-10-138-473-34
5	214	15.2	297	5	US-10-450-726-4
6	208	14.8	295	4	US-10-156-761-8839
7	207	14.7	301	3	US-09-738-626-6745
8	207	14.7	301	3	US-09-746-660A-20
9	207	14.7	301	3	US-09-746-660A-20
10	207	14.7	301	3	US-10-450-055-8
11	207	14.7	301	6	US-11-006-098-314
12	206.5	14.6	290	4	US-10-156-761-8814
13	180	12.8	300	4	US-10-282-122A-53861
14	131	9.3	181	4	US-10-437-963-126675
15	127.5	9.0	3227	5	US-10-760-493-33
16	118.5	8.4	2539	4	US-10-282-122A-61655
17	113	8.0	7068	4	US-10-203-295-20
18	113	8.0	9477	4	US-10-203-295-37
19	112	7.9	11088	4	US-10-203-295-7
20	112	7.9	11096	5	US-10-732-923-20557
21	109.5	7.8	10917	5	US-10-732-923-20606
22	109	7.7	2365	4	US-10-156-761-14890
23	108	7.7	1050	4	US-10-156-761-8816
24	107.5	7.6	801	4	US-10-368-493-8838
25	106.5	7.6	860	5	US-10-925-357-22
26	106	7.5	1641	3	US-09-980-217-30
27	106	7.5	3753	3	US-09-980-217-29

28	105.5	7.5	152	4	US-10-282-122A-45977	Sequence 45977, A
29	105.5	7.5	471	4	US-10-437-963-156094	Sequence 156094, A
30	105.5	7.5	1976	5	US-10-937-730A-7	Sequence 7, Appl1
31	105.5	7.5	6291	4	US-10-329-079-41	Sequence 41, Appl1
32	103.5	7.3	533	4	US-10-156-761-13934	Sequence 13934, A
33	103	7.3	181	4	US-10-425-115-230179	Sequence 230179, A
34	103	7.3	409	4	US-10-156-761-13624	Sequence 13624, A
35	103	7.3	3025	3	US-09-980-217-19	Sequence 19, Appl1
36	103	7.3	5435	5	US-10-203-295-38	Sequence 38, Appl1
37	103	7.3	5435	5	US-10-732-923-20558	Sequence 20558, A
38	103	7.3	7429	5	US-10-937-730A-5	Sequence 5, Appl1
39	102.5	7.3	1657	5	US-10-760-493-27	Sequence 27, Appl1
40	102.5	7.3	10625	5	US-10-819-386A-3	Sequence 3, Appl1
41	102	7.2	890	5	US-10-732-923-8067	Sequence 8067, Ap
42	101.5	7.2	391	4	US-10-344-738-31	Sequence 31, Appl1
43	101.5	7.2	402	4	US-10-369-493-22935	Sequence 22935, A
44	101.5	7.2	795	4	US-10-314-657-43	Sequence 43, Appl1
45	101.5	7.2	795	5	US-10-473-193-43	Sequence 43, Appl1

ALIGNMENTS

RESULT 1
US-10-617-038-21
; Sequence 21, Application US/10617038
; Publication No. US20040057963A1
; GENERAL INFORMATION:
; APPLICANT: Andersen, Peter
; APPLICANT: Rosenkrands, Ida
; APPLICANT: Strjhn, Anette
; TITLE OF INVENTION: Therapeutic TB Vaccine
; FILE REFERENCE: SSI5AUSA
; CURRENT APPLICATION NUMBER: US/10/617, 038
; CURRENT FILING DATE: 2003-07-11
; PRIOR APPLICATION NUMBER: DK PA 2002 01098
; PRIOR FILING DATE: 2002-07-13
; PRIOR APPLICATION NUMBER: US 60/401, 725
; PRIOR FILING DATE: 2002-08-07
; NUMBER OF SEQ ID NOS: 187
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 21
; LENGTH: 279
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-617-038-21

Query Match 100.0%; Score 1410; DB 4; Length 279;
Best Local Similarity 100.0%; Pred. No. 4.8e-132;
Matches 279; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MNOSHKPESTIVGIDGSKPAVOALAVDEAASRDILRLLYAIEPDPGVAHGAARK	60
DB	1	MNOSHKPESTIVGIDGSKPAVOALAVDEAASRDILRLLYAIEPDPGVAHGAARK	60
QY	61	LAALANVRAFTVAEADRPVKEVEITQERPTSLIRASAAALVCVGAIGHHRRPE	120
DB	61	LAALANVRAFTVAEADRPVKEVEITQERPTSLIRASAAALVCVGAIGHHRRPE	120
QY	121	RVGSTAALALASQCPVAIVRPHRPVIGRDAAMIVVEADSSDVGVLGAVMAEARLDS	180
DB	121	RVGSTAALALASQCPVAIVRPHRPVIGRDAAMIVVEADSSDVGVLGAVMAEARLDS	180
QY	181	PVRVVTOROGVGTGDVVASLDRWLARWQPRPDVRSAAVHGBLLDYLGLGSRVH	240
DB	181	PVRVVTOROGVGTGDVVASLDRWLARWQPRPDVRSAAVHGBLLDYLGLGSRVH	240
QY	241	MVVLASDQEHVEQLVGPAGNAVLQEAAGCTLLVVGQOYL	279
DB	241	MVVLASDQEHVEQLVGPAGNAVLQEAAGCTLLVVGQOYL	279

RESULT 2

```

Query Match          30.6%; Score 431.5; DB 4; Length 268;
Best Local Similarity 40.4%; Pred. No. 2,2e-34;
Matches 111; Conservative 34; Mismatches 105; Indels 25; Gaps 8;

QY      1  MNQGHKPSIVIGIDGSKPAVOALIMAWDEAASRDIPRLYLAIPEDDPGYAAGMAARK 60
DB      1  MSDRPAPAAVVVIGIDGSRPAATHAALMAVDEAVNRDIPRLVYVIDPSQLSAAGEGCG-- 58
QY      61  LAALENAVRYAFTVAEADRPVKVEVEITTOERPVYSLIRASAAALVCGAIGVHHFRPE 120
DB      59  -SAAPAAALHDHSRKKEATGQEPKLTETELCGPRLTKYMQESRSAAMLCVGSGVDHVGRR 117
QY      121  RVGSTAAALALASQCPVAIVRPH--RVPIDGRDAIVVEADGSDIGVLGAVMAEARLR 178
DB      118  R-GSVAATLACASALCPVAIVHPSPAERPTTSQVSAVAEVRN----GVLVHAPEEARLR 172
QY      179  DSPRVVYTCROSGVGDTGDDV-----RASLDKWLAFMOPRPYDPVRVQSAAVHGGELIDY 231
DB      173  GVPLRAVAVH---AAETPDDEQGSRLAHVHLSRRLAMTRLYPEVRVDRAIAGGSACRH 229
QY      232  LAGLGRSVHMYVYLASDQEHVEQLVGA--PGNAVY 264
DB      230  LAANAKPEQLFV---ADSHSAHELCGAIQPGCAVL 261

RESULT 3
US-10-617-038-27
; Sequence 27, Application US/10617038
; Publication No. US20040057963A1
; GENERAL INFORMATION:
;   APPLICANT: Andersen, Peter
;   APPLICANT: Rosenkrands, Ida
;   APPLICANT: Stryhn, Anette
;   TITLE OF INVENTION: Therapeutic TB Vaccine
;   FILE REFERENCE: SSI5AUSA
;   CURRENT APPLICATION NUMBER: US/10/617,038
;   PRIOR FILING DATE: 2003-07-11
;   PRIOR APPLICATION NUMBER: DK PA 2002 01098
;   PRIOR APPLICATION NUMBER: US 60/401,725
;   PRIOR FILING DATE: 2002-08-07
;   NUMBER OF SEQ ID NOS: 187
;   SOFTWARE: PatentIn version 3.2
;   SEQ ID NO 27
;   LENGTH: 272
;   TYPE: PRT
;   ORGANISM: Mycobacterium tuberculosis
US-10-617-038-27
Query Match          29.5%; Score 416; DB 4; Length 272;

```

```

      RESULT 4
US-10-138-473--34
; Sequence 34, Application US/10138473
; Publication No. US20030165525A1
GENERAL INFORMATION:
APPLICANT: ANDERSEN, Peter
APPLICANT: WELDSINGH, Karin
APPLICANT: HANSEN, Christina Veggerby
APPLICANT: FLORIO, Walter
APPLICANT: OKKELS, Li Wei Meng
APPLICANT: SKJORT, Rikke Louise Vinther
APPLICANT: ROSENKRANDS, Ida
APPLICANT: RASMUSSEN, Peter Birk
TITLE OR INVENTION: TB diagnostic Based On Antigens From M. tuberculosis
FILE REFERENCE: 0459-0710P
CURRENT APPLICATION NUMBER: US/10/138,473
CURRENT FILING DATE: 2002-05-02
PRIOR APPLICATION NUMBER: US 10/060,428
PRIOR FILING DATE: 2002-01-29
PRIOR APPLICATION NUMBER: US 09/415,884
PRIOR FILING DATE: 1999-10-08
PRIOR APPLICATION NUMBER: US 60/116,673
PRIOR FILING DATE: 1999-01-21
PRIOR APPLICATION NUMBER: DK 1998 01281
PRIOR FILING DATE: 1998-10-18
PRIOR APPLICATION NUMBER: US 60/070,488
PRIOR FILING DATE: 1998-01-05
PRIOR APPLICATION NUMBER: DK 1997 01277
PRIOR FILING DATE: 1997-11-10
PRIOR APPLICATION NUMBER: US 60/044,624
PRIOR FILING DATE: 1997-04-18
PRIOR APPLICATION NUMBER: DK 1997 00376
PRIOR FILING DATE: 1997-04-02
NUMBER OF SEQ ID NOS: 174
SOFTWARE: PatentIn version 3.1
SEQ ID NO 34
LENGTH: 297
TYPE: PRT
ORGANISM: Mycobacterium tuberculosis
US-10-138-473--34

Query Match          15.2%; Score 214; DB 4; Length 297;
Best Local Similarity 27.2%; Pred. No. 1,3e-12;
Matches 84; Conservative 37; Mismatches 134; Indels 54; Gaps 10;

Cy      1 MNOSHKRPSIVGVGIDGSKPAVOALWVDEASRDIPRLRLVYAIEPD-----DPGY 51
        | : ||| | | | | | | | | | | | | | | | | | | | | | | | |
Db      1 MSSNSLSLGIITVGIDSPAAQAVVRNARADAEIRKILPLTVHAVSPPVATWLEVPPLPGV 60

```


QY 52 AA-----HGAARKLAAMENAVTAFTVEAD---RPVKVEVEITQERPTVSLIRASAAA 104
DB 61 LRWQODHG-----RHLLDDALKVVEQASLRAGPETHSEITVPAAVPTLVDMSDA 111
QY 105 ALVVCAGIAGVHFRPERVSTAAALALSAQCPVAIV-----RPHRVPIGRDAAMIV 156
DB 112 VLMVYVCGCLSGSRWGRRLGVSGLLRHAKCPVYIITHEDSVMHFQCAP-----VLV 164
QY 157 EADGSSDIGVLLGAVMAEARLRDSPRVVTCRGSGVGT-----GDDVRAAS-----LDR 205
DB 165 GVDGSSASSELATATAIFDEASRRN--VDLVALHMSDVDSWPGIDMPATQSMALQVLA 222
QY 206 WLAAMQRRYPRVYQSAAVHGEELDYLAGLGRSVHMYVLSASDQ--EHVEQLVGAPGNAV 264
DB 223 RLAWQMRERYPVATVTRVVVRDQPARQLVQRSEEAQLVVGSGRGYAGMLVGSVGTVA 282
QY 265 QEAGCTLLV 273
DB 283 QLAATPVIV 291

RESULT 5

US-10-450-726-4
; Sequence 4, Application US/10450726
; Publication No. US20040242471A1
; GENERAL INFORMATION:
; APPLICANT: Dick, Thomas
; APPLICANT: Boon, Calvin Ka Khin
; TITLE OF INVENTION: DORMANCY-INDUCED MYCOBACTERIUM PROTEINS
; FILE REFERENCE: 50316/002001
; CURRENT APPLICATION NUMBER: US/10/450, 726
; CURRENT FILING DATE: 2003-06-13
; PRIOR APPLICATION NUMBER: PCT/EP01/14551
; PRIOR FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: 0030368.5
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 297
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-450-726-4

Query Match 15.2%; Score 214; DB 5; Length 297;
Best Local Similarity 27.2%; Pred. No. 1.3e-12;
Matches 84; Conservative 37; Mismatches 134; Indels 54; Gaps 10;

QY 1 MNOSHKPPSIYVIGIDSKPAVOALMAVDEASFDIPRLLYAIEPD-----DPGY 51
DB 1 MSSGNSSLGIIVGIDSDPAQVAVRMAARDAELRKPLTLVHAVSPVATWLEVPPLPGV 60
QY 52 AA-----HGAARKLAAMENAVTAFTVEAD---RPVKVEVEITQERPTVSLIRASAAA 104
DB 61 LRWQODHG-----RHLLDDALKVVEQASLRAGPETHSEITVPAAVPTLVDMSDA 111
QY 105 ALVVCAGIAGVHFRPERVSTAAALALSAQCPVAIV-----RPHRVPIGRDAAMIV 156
DB 112 VLMVYVCGCLSGSRWGRRLGVSGLLRHAKCPVYIITHEDSVMHFQCAP-----VLV 164
QY 157 EADGSSDIGVLLGAVMAEARLRDSPRVVTCRGSGVGT-----GDDVRAAS-----LDR 205
DB 165 GVDGSSASSELATATAIFDEASRRN--VDLVALHMSDVDSWPGIDMPATQSMALQVLA 222
QY 206 WLAAMQRRYPRVYQSAAVHGEELDYLAGLGRSVHMYVLSASDQ--EHVEQLVGAPGNAV 264
DB 223 RLAWQMRERYPVATVTRVVVRDQPARQLVQRSEEAQLVVGSGRGYAGMLVGSVGTVA 282
QY 265 QEAGCTLLV 273
DB 283 QLAATPVIV 291

RESULT 6

US-10-156-761-8839
; Sequence 8839, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156, 761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 8839
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-8839

Query Match 14.8%; Score 208; DB 4; Length 295;
Best Local Similarity 28.2%; Pred. No. 5.2e-12;
Matches 87; Conservative 39; Mismatches 117; Indels 66; Gaps 11;

QY 9 SIYVIGIDSKRPVQALMAVDEASRDIPLRLLYAIEP-----DDPGAAHGA 58
DB 7 TIVGIDGSRGTGDAADWAAREARRRLPLRLHAGAEPPVPAVPVDDVLPARTALDR 66
QY 59 RLKLAAMENAVTAFTVEADRPVKVEVEITQERPTVSLIRASAAAALVVCAGIAGVHFR 118
DB 67 VOLA-----YHAPALDIARR-----TETPAVALAAMAEATLVLSRGTFGA 112
QY 119 PERVSTAAALALSAQCPVAIVRPHRV-----PIGRDAAM--IIVENDGSSDI 164
DB 113 GFLVGSVALAVMAARARPVVLTGTGLAEDERPADGCTPGR--AAVLPVVLGLDLHRA 171
QY 165 GVLGAVMAEARLRDSPRVV-----TCRGSGVGTGD-----DVASLDRWLAR 210
DB 172 DRLLAYAFBAAVRASAPLVHVTWTLPPGHAAPGAPLPEDAAREBGARTLTLTLPW 231
QY 211 QPRYPDVYQSAAVHGEELDYLAGLGRSVHMYVLSASDQEHVEQLVGAPG-----NAV 265
DB 232 RHKFPGTDLERYIPGHGHLRLASIRAGLVIG-----RTSAPGIGRAARSLIH 283
QY 266 EAGCTLLV 274
DB 284 HAGCPVAV 292

RESULT 7

US-09-738-626-6745
; Sequence 6745, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHITO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125

```

; NUMBER OF SEQ ID NOS: 125
; SOFTWARE: PatentIn Vers. 2.0
; SEQ ID NO 6
; LENGTH: 301
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-746-660A-6

Query Match      14.7%; Score 207; DB 3; Length 301;
Beet Local Similarity 26.1%; Pred. No. 6,7e-12;
Matches 77; Conservative 46; Mismatches 134; Indels 38; Gaps 10;

QY      10  IVGVIGDSKKPVAQALNAVDEAASRDIPLRL-----LYA---IEPDDPGVAANGA 57
      11  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      7  VVAVDGSDASKQVRMAAANTANRRGILPLRIASSYTPQFLYEGWVPPELEDDLOAEA 66
      8  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      58  ARKAAAEENARVAFYTVAAADRPVKYVEITQERPVTSILRASAAALVCVGAIGVHHF 117
      59  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      67  LEKINEARD-----IAHEVAPE-IKGHITABSPFDIMLLNSPDATIMVMSGRIGSL 119
      68  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      118  RPERVGSYTAALALSAQCPSVAIYVPHRVPYIGRDAW--IYVEADGSSDGIUIGAVMAEA 175
      119  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      120  SCGMWGSVGAIVSHAKCPVVVAREDSA-VNEDSKYGPVVVGVDSEVSQATEYAFABA 178
      121  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      176  RLRF-----DSFRRVYTCQSGSVGTGDDY-RASLDRLAKRQF--RYPRVRY 220
      177  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      179  EARGAEIYAVHTYMDQVQASLAGLAAQQOWEVEVQOTDMLIERLAPLVERKPSYTVK 238
      180  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      221  SAAVHGEUUDYIAGLRGVHVVVLSASDOHVE-QLVGAPGNVYLOEAGCTLLV 274
      222  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      239  KIITRDRPVRLALAEASENAQLLVGSHGRGFGKMLIGSTRALLQSAFCPMNV 293
      240  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 9
US-09-746-660A-20
; Sequence 20, Application US/09746660A
; Publication No. US20030049804A1
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Krieger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zeider, Oskar
; APPLICANT: Haberhauser, Gregor
; APPLICANT: Kim, Jun-Won
; APPLICANT: Lee, Heung-Schick
; APPLICANT: Hwang, Byung-Joon
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING
; FILE REFERENCE: BGI-121CP2
; CURRENT APPLICATION NUMBER: US/09/746,660A
; CURRENT FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 09/606740
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 09/603124
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 60/142101
; PRIOR FILING DATE: 1999-07-02
; PRIOR APPLICATION NUMBER: 60/148613
; PRIOR FILING DATE: 1999-08-12
; PRIOR APPLICATION NUMBER: 60/187970
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: DE 19931420.9
; PRIOR FILING DATE: 1999-07-08
; NUMBER OF SEQ ID NOS: 125
; SOFTWARE: PatentIn Vers. 2.0
; SEQ ID NO 20
; LENGTH: 301
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-746-660A-20

```

```

Query Match      14.7%; Score 207; DB 3; Length 301;
Best Local Similarity 26.1%; Pred. No. 6,7e-12;
Matches 77; Conservative 46; Mismatches 134; Indels 38; Gaps 10;

Qy 10 IVGIDGSKPAVOALMAVDEAASRDIPRL-----LYA---IEPDDGYAAHGAA 57
   :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 7 VVVAVDGSDASKQAVRMAANTANKRGIPRLASSYTMPQFLYAEGWVPPOELFDLQAEA 66

Qy 58 ARKLAASNAVRVAFVAVEAADRPVKVEVEITQERPTSLIRASAAALVCYGAIGHNF 117
   :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 67 LEKINEAD-----IAHEVAPE-IKIGHTIABGSPIDMLEMSDATTIMVSGRGLGL 119

Qy 118 RPERVSTAAALALASQCPVAIVRPHRVPFIGRDAW--IYVEADGSSDIGVLGAWMAEA 175
   :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 120 SGWVGVSAGAVVSHAKCPVVVVRSDA-VNEDSKYGPVVVGVDGSEVSQATEYAFAEA 178

Qy 176 RLR-----DSPRVVTCRQSGVGTGDDV-RASLDRLMARQP---RYPDVAVQ 220
   :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 179 EARGAEIVAVHTWMDMQVQASLAGLAAAOQOWDEVERQOTDMLERLAPLVEKYPSTVVK 238

Qy 221 SAAVHGBLDYLAGRSVHMVVLASDOEHVE-QLVGAPGNVQLQAGCTLLVY 274
   :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 239 KIITRDRPVRLAEASNAQLLVGSHRGGFKMGLDSTSRALLQSAPCPMNV 293

RESULT 10
US-10-450-055-8
; Sequence 8, Application US/10450055
; Publication No. US20040043953A1
; GENERAL INFORMATION:
; APPLICANT: BASF Aktiengesellschaft
; TITLE OF INVENTION: No. US20040043953A1el genes of Corynebacterium
; FILE REFERENCE: 936 2000
; CURRENT FILING DATE: 2003-06-10
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Vers. 2.0
; SEQ ID NO 8
; LENGTH: 301
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-10-450-055-8

Query Match      14.7%; Score 207; DB 4; Length 301;
Best Local Similarity 26.1%; Pred. No. 6,7e-12;
Matches 77; Conservative 46; Mismatches 134; Indels 38; Gaps 10;

Qy 10 IVGIDGSKPAVOALMAVDEAASRDIPRL-----LYA---IEPDDGYAAHGAA 57
   :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 7 VVVAVDGSDASKQAVRMAANTANKRGIPRLASSYTMPQFLYAEGWVPPOELFDLQAEA 66

Qy 58 ARKLAASNAVRVAFVAVEAADRPVKVEVEITQERPTSLIRASAAALVCYGAIGHNF 117
   :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 67 LEKINEAD-----IAHEVAPE-IKIGHTIABGSPIDMLEMSDATTIMVSGRGLGL 119

Qy 118 RPERVSTAAALALASQCPVAIVRPHRVPFIGRDAW--IYVEADGSSDIGVLGAWMAEA 175
   :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 120 SGWVGVSAGAVVSHAKCPVVVVRSDA-VNEDSKYGPVVVGVDGSEVSQATEYAFAEA 178

Qy 176 RLR-----DSPRVVTCRQSGVGTGDDV-RASLDRLMARQP---RYPDVAVQ 220
   :|:|:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~
Db 179 EARGAEIVAVHTWMDMQVQASLAGLAAAOQOWDEVERQOTDMLERLAPLVEKYPSTVVK 238

Qy 221 SAAVHGBLDYLAGRSVHMVVLASDOEHVE-QLVGAPGNVQLQAGCTLLVY 274
   :|:|:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~
Db 239 KIITRDRPVRLAEASNAQLLVGSHRGGFKMGLDSTSRALLQSAPCPMNV 293

RESULT 11
US-11-006-098-314
; Sequence 314, Application US/11006098
; Publication No. US20050153402A1
; GENERAL INFORMATION:

```

```

; APPLICANT: Pompeius, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zeider, Oskar
; APPLICANT: Haberhauer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING
; FILE REFERENCE: BGI-123CP
; CURRENT FILING DATE: US/11/006,098
; CURRENT FILING DATE: 2004-12-06
; PRIOR APPLICATION NUMBER: US/09/602,874
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 60/142690
; PRIOR FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: 60/151251
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19930476.9
; PRIOR FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: DE 19931419.5
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: DE 19931420.9
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19932122.1
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932128
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932206.6
; PRIOR FILING DATE: 1999-07-09
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 362
; SEQ ID NO 314
; LENGTH: 301
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-11-006-098-314

Query Match      14.7%; Score 207; DB 6; Length 301;
Best Local Similarity 26.1%; Pred. No. 6,7e-12;
Matches 77; Conservative 46; Mismatches 134; Indels 38; Gaps 10;

Qy 10 IVGIDGSKPAVOALMAVDEAASRDIPRL-----LYA---IEPDDGYAAHGAA 57
   :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:~:~:~
Db 7 VVVAVDGSDASKQAVRMAANTANKRGIPRLASSYTMPQFLYAEGWVPPOELFDLQAEA 66

Qy 58 ARKLAASNAVRVAFVAVEAADRPVKVEVEITQERPTSLIRASAAALVCYGAIGHNF 117
   :|:|:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~
Db 67 LEKINEAD-----IAHEVAPE-IKIGHTIABGSPIDMLEMSDATTIMVSGRGLGL 119

Qy 118 RPERVSTAAALALASQCPVAIVRPHRVPFIGRDAW--IYVEADGSSDIGVLGAWMAEA 175
   :|:|:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~
Db 120 SGWVGVSAGAVVSHAKCPVVVVRSDA-VNEDSKYGPVVVGVDGSEVSQATEYAFAEA 178

Qy 176 RLR-----DSPRVVTCRQSGVGTGDDV-RASLDRLMARQP---RYPDVAVQ 220
   :|:|:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~
Db 179 EARGAEIVAVHTWMDMQVQASLAGLAAAOQOWDEVERQOTDMLERLAPLVEKYPSTVVK 238

Qy 221 SAAVHGBLDYLAGRSVHMVVLASDOEHVE-QLVGAPGNVQLQAGCTLLVY 274
   :|:|:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~:~
Db 239 KIITRDRPVRLAEASNAQLLVGSHRGGFKMGLDSTSRALLQSAPCPMNV 293

RESULT 12
US-10-156-761-8814
; Sequence 8814, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI

```

APPLICANT: SAKAKI, YOSHIYUKI
 APPLICANT: HATTORI, MASAHIRA
 TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 FILE REFERENCE: 249-262
 CURRENT APPLICATION NUMBER: US/10/156,761
 CURRENT FILING DATE: 2002-05-29
 PRIOR APPLICATION NUMBER: JP 2001-204089
 PRIOR FILING DATE: 2001-05-30
 PRIOR APPLICATION NUMBER: JP 2001-272697
 PRIOR FILING DATE: 2001-08-02
 NUMBER OF SEQ ID NOS: 15109
 SEQ ID NO 8814
 LENGTH: 290
 TYPE: PRT
 ORGANISM: Streptomyces avermitilis
 US-10-156-761-8814

Query Match 14.6%; Score 206.5; DB 4; Length 290;
 Best Local Similarity 26.9%; Pred. No. 7.2e-12;
 Matches 77; Conservative 46; Mismatches 138; Indels 25; Gaps 8;

QY 10 IVVGIDSKPAVQALMAVDEAASRDIPRLVYI--EPDDPGYAAHG-AAARKLAAEN 66
 DB 5 LVAVDSESSMRVADMAADEALRGVPLRLVYLSLMEYRGALADGLGRPEQVYRDD 64
 QY 67 AVRVAFTAVEAADRPVKVEVEITQERPVTSILIRASAAALVCVGAIGVHFRPERVSTA 126
 DB 65 IVEAAGRAHRRHADVKISTDVLPEEPAPALLREGNEAFALVLSGRSGLAELLGVS 124
 QY 127 AALLSACCPVAIYR--PHRVPIGRDAWIIVEADGSSDIGVLGAVMAARLDSYR 183
 DB 125 LAVAARAYCPVYIVRGSDNRGAGHRRIVLAVGEBDDPAARFAV-ABARARGCAVE 183
 QY 184 VVTC---ROSGVDPTGDDVRAS-----LDRMLARWOPKRPDVRVOSAIVHGL 229
 DB 184 AVRWRPARRALGHPLSSRRRAARHERATEVLEAVS---GAPDVELRRRAVEGPAR 240
 QY 230 DYLAGRSVHVYVLSASDOE-HVEQIVGABGNVLOEAGCTLLV 274
 DB 241 KVLDDASATADLLVGARRHGHSGQLQGRVAAHMLHHSACPAVV 286

RESULT 13
 US-10-282-122A-53861
 Sequence 53861, Application US/10282122A
 Publication No. US20040029129A1
 GENERAL INFORMATION:
 APPLICANT: Wang, Liangou
 APPLICANT: Zamudio, Carlos
 APPLICANT: Malone, Cheryl
 APPLICANT: Haselebeck, Robert
 APPLICANT: Ohlsen, Kari
 APPLICANT: Zykend, Judith
 APPLICANT: Wall, Daniel
 APPLICANT: Trawick, John
 APPLICANT: Carr, Grant
 APPLICANT: Yamamoto, Robert
 APPLICANT: Forsyth, R.
 APPLICANT: Xu, H.
 TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
 FILE REFERENCE: ELITRA.034A
 CURRENT APPLICATION NUMBER: US/10/282,122A
 CURRENT FILING DATE: 2003-02-20
 PRIOR APPLICATION NUMBER: 60/191,078
 PRIOR FILING DATE: 2000-03-21
 PRIOR APPLICATION NUMBER: 60/206,848
 PRIOR FILING DATE: 2000-05-23
 PRIOR APPLICATION NUMBER: 60/207,727
 PRIOR FILING DATE: 2000-05-26
 PRIOR APPLICATION NUMBER: 60/230,335
 PRIOR FILING DATE: 2000-09-06
 PRIOR APPLICATION NUMBER: 60/230,347
 PRIOR FILING DATE: 2000-09-09

PRIOR APPLICATION NUMBER: 60/242,578
 PRIOR FILING DATE: 2000-10-23
 PRIOR APPLICATION NUMBER: 60/253,625
 PRIOR FILING DATE: 2000-11-27
 PRIOR APPLICATION NUMBER: 60/257,931
 PRIOR FILING DATE: 2000-12-22
 PRIOR APPLICATION NUMBER: 60/267,636
 PRIOR FILING DATE: 2001-02-09
 PRIOR APPLICATION NUMBER: 60/269,308
 PRIOR FILING DATE: 2001-02-16
 Remaining Prior Application data removed - See file wrapper or PALM.
 NUMBER OF SEQ ID NOS: 78614
 SOFTWARE: Patentin version 3.1
 SEQ ID NO 53861
 LENGTH: 300
 TYPE: PRT
 ORGANISM: Corynebacterium diptheriae
 US-10-282-122A-53861

Query Match 12.8%; Score 180; DB 4; Length 300;
 Best Local Similarity 24.8%; Pred. No. 3.3e-09;
 Matches 73; Conservative 46; Mismatches 139; Indels 36; Gaps 9;

QY 10 IVVGIDSKPAVQALMAVDEAASRDIPRL-----LYA---IEPDDPGYAAHGAA 57
 DB 7 VVAVDSSEASQNAVRAANTANKRGVPLRLAASVTPQFLVAGVWPQELFDELQSET 66
 QY 58 ARKLAAMENVRVAFTVEAADRPVKVEVEITQERPVTSILIRASAAALVCVGAIGVHFR 117
 DB 67 MDMEIAR-----VVAHEVAP-DIKIGVYIAGSPIDMLDMSDVDTMIVMSRGIGL 119
 QY 118 RPERVSTAALALSACCPVAIYR-PHRVPIGRDAWIIVEADGSSDIGVLGAVMAEAR 176
 DB 120 SGWVGVSAAVYSHADCPVVVVRSDVHTETNKYGPVVVGVDVSGRATFAPFEBAQ 179
 QY 177 LRDSR-VRVVTCQSGVDGDDVRASLDW-----LARMP---RPDVRVOS 221
 DB 180 ARGKLVAIHTMDMQVQASLAGLAAQOEWEIIEKQTTLLKDRQPLLERPDEVEEM 229
 QY 222 AAHVHGLDVIAGRSVHVYVLSASDOE-HVEQIVGABGNVLOEAGCTLLV 274
 DB 240 VTRDRPVRLBEDCAHNAQLLVVSGHSGFRGMLGSTRALLQSNAPCMVV 293

RESULT 14
 US-10-437-963-126675
 Sequence 126675, Application US/10437963
 Publication No. US20040123343A1
 GENERAL INFORMATION:
 APPLICANT: La Rosa, Thomas J.
 APPLICANT: Kovalic, David K.
 APPLICANT: Zhou, Yihua
 APPLICANT: Cao, Yongwei
 APPLICANT: Wu, Wei
 APPLICANT: Boukharov, Andrey A.
 APPLICANT: Barbazuk, Brad
 APPLICANT: Li, Ping
 TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
 FILE REFERENCE: 38-21(53221)B
 CURRENT APPLICATION NUMBER: US/10/437,963
 CURRENT FILING DATE: 2003-05-14
 NUMBER OF SEQ ID NOS: 204966
 SEQ ID NO 126675
 LENGTH: 181
 TYPE: PRT
 ORGANISM: Oryza sativa
 FEATURE:
 OTHER INFORMATION: Clone ID: PAT_MRT4530_29199C.1.pep
 US-10-437-963-126675

Query Match 9.3%; Score 131; DB 4; Length 181;
 Best Local Similarity 30.4%; Pred. No. 0.00013;

Matches 51; Conservative 23; Mismatches 54; Indels 40; Gaps 7;

Job time : 60.6719 secs

QY 10 IVIGIGGSKAAVAALAAVAD-----EASRDIP--RLIYALE-----P 46
 Db 3 VLVAVDDSKSRHSLSWLDHLFFPPLAATGDCGEEQVPRPAPELVYHALLEPLHYMFP 62
 QY 47 DDEGYAAHGAA-----ARKLAAENAVRAFTAVEADR-----PVKVEVEITQERPTS 96
 Db 63 VEGGSNAVYGAASMMELVRAAQAEMNANLLVRKLLCERGVAAALVAVEGE-----PREA 117
 QY 97 LIRAS--AAALATVCGAIGVHHFRPERVGTAAALATASOCPVALVR 142
 Db 118 LCPAAEDAGAGLLVVGSRGLAKKAFVLSGSVDYCAHNASCPIMVKP 165

RESULT 15

```

US-10-760-493-33
; Sequence 33, Application US/10760493
; Publication No. US20050187167A1
;
; GENERAL INFORMATION:
; APPLICANT: Ecopia Biosciences Inc
; APPLICANT: Farnet, Chris M.
; APPLICANT: McAlpine, James B.
; APPLICANT: Zazopoulos, Emmanuel
; APPLICANT: Bachmann, Brian O.
; APPLICANT: Pitraee, Mahmood
; TITLE OF INVENTION: POLYENE POLYKETIDES, PROCESSES FOR THEIR PRODUCTION AND THEIR USE
; FILE REFERENCE: 3004-9US
; CURRENT APPLICATION NUMBER: US/10/760,493
; CURRENT FILING DATE: 2004-01-21
; PRIOR APPLICATION NUMBER: US5N 60/441,123
; PRIOR FILING DATE: 2003-01-21
; PRIOR APPLICATION NUMBER: US5N 60/469,810
; PRIOR FILING DATE: 2003-05-13
; PRIOR APPLICATION NUMBER: US5N 60/491,516
; PRIOR FILING DATE: 2003-08-01
; PRIOR APPLICATION NUMBER: US5N 60/494,568
; PRIOR FILING DATE: 2003-08-13
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 33
; LENGTH: 3227
; TYPE: PR1
; ORGANISM: Streptomyces aizunensis
; US-10-760-493-33

```

Query Match	9.0%;	Score 127.5;	DB 5;	Length 3227;
Best Local Similarity	27.6%;	Pred. No. 0.015;		
Matches	81;	Conservative	23;	Mismatches 115;
			Indels	75;
			Gaps	14;

QY	41	YALIEPDDPGVAHGAARXKLAAAMENAYFAFTVEADR	-----PVAVE-----	85	
Db	2592	LVVPENGPDEMTGAVLRVLADRGAEVRYTVVPADGCDRRLAATLAETDGAAPAGVL		2655	
QY	86	-----VEITGERPVTSLIRASA-----	AAALVCV--GAIGVHFRERVGSTAA	127	
Db	2652	SLALAVESALERTHTGTLATATAALQALGDADVAPALCCTRGAVSV--	ARTERLQDPAQ	2709	
QY	128	ALALSAQCPVAVLPHR-----VPIGHDAMI-----	YVADGSSD-----IGVLGAV	171	
Db	2710	ALVSGFGRTVLALEYPRDWGGLVLDLPEQADGRTLELACVLADGSGEDVALRASGLFERR		2765	
QY	172	MAEARLRDSPVRVVTQROSG-----	VGDYGDVVRASLIDRWLAR-----	WQRYPDVRYV	219
Db	2770	LVAHPLADLT--AAVREMRPQGTTLVYGGTG--	ALGHAVARMLAENGAHEHLILTSRRGPD---	2824	
QY	220	QSAAVHGGILLVYLACLGERSYHMYVULSADQDERVEQVQA-----	PGNAVLDQAG	268	
Db	2825	--AGAAELRELTLGAQVTTATCDMAKDRDVAALIAAVPADQCLTVYVHMTAG		2876	

Search completed: March 23, 2006, 06:28:09

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: March 23, 2006, 06:17:23 ; Search time 6.6953 Seconds

(without alignments)
1192.746 Million cell updates/sec

Title: US-10-617-038-21

Perfect score: 1410
Sequence: 1 MNSGKPSIIVGIDGSKPA.....GNAVLOEAGCTLVVGOQYL 279

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 169630 seqs, 28622889 residues

Total number of hits satisfying chosen parameters: 169630

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_AA_New:*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	209	14.8	305	US-11-087-099-3468	Sequence 3468, Ap
2	207	14.7	301	US-11-055-822-14	Sequence 14, Appl
3	136	9.6	152	US-11-087-099-6347	Sequence 6347, Ap
4	99.5	7.1	828	US-11-087-099-6863	Sequence 6863, Ap
5	98.5	7.0	171	US-11-096-568A-32572	Sequence 32572, A
6	98.5	7.0	179	US-11-096-568A-32572	Sequence 32572, A
7	97	6.9	357	US-11-096-568A-17171	Sequence 17171, A
8	97	6.9	370	US-11-096-568A-17170	Sequence 17170, A
9	95.5	6.8	316	US-11-096-568A-13340	Sequence 13340, A
10	94.5	6.7	345	US-11-096-568A-18547	Sequence 18547, A
11	93.5	6.6	161	US-11-087-099-2546	Sequence 2546, Ap
12	93.5	6.6	161	US-11-087-099-6989	Sequence 6989, Ap
13	93.5	6.6	333	US-11-096-568A-18548	Sequence 18548, A
14	91	6.5	174	US-11-096-568A-27471	Sequence 27471, A
15	90.5	6.4	175	US-10-454-437-254	Sequence 254, Ap
16	90.5	6.4	1086	US-11-052-554A-380	Sequence 380, Ap
17	90.5	6.4	3689	US-11-075-185-4	Sequence 185, Appl
18	90	6.4	592	US-10-840-688-22	Sequence 22, Appl
19	90	6.4	7968	US-11-143-980-49	Sequence 49, Appl
20	90	6.4	8695	US-11-205-109-15	Sequence 15, Appl
21	89	6.3	268	US-11-096-568A-18549	Sequence 18549, A
22	88.5	6.3	159	US-11-087-099-12270	Sequence 12270, A
23	88	6.2	7102	US-11-143-980-48	Sequence 48, Appl
24	87	6.2	150	US-11-087-099-3411	Sequence 3411, Ap
25	87	6.2	162	US-11-087-099-5731	Sequence 5731, Ap

26	86.5	6.1	920	US-11-087-099-11404	Sequence 11404, A
27	86	6.1	152	US-11-096-568A-4792	Sequence 4792, Ap
28	86	6.1	163	US-11-087-099-11409	Sequence 11409, A
29	86	6.1	163	US-11-096-568A-4791	Sequence 4791, Ap
30	86	6.1	164	US-11-087-099-6603	Sequence 6603, Ap
31	86	6.1	165	US-11-096-568A-4790	Sequence 4790, Ap
32	85.5	6.1	181	US-11-087-099-10726	Sequence 10726, A
33	85.5	6.1	852	US-10-645-441-15	Sequence 15, Appl
34	85.5	6.1	852	US-10-725-475-7	Sequence 7, Appl
35	85.5	6.1	852	US-11-050-804-6	Sequence 6, Appl
36	85.5	6.1	3300	US-11-052-554A-133	Sequence 133, Ap
37	85	6.0	316	US-11-096-568A-16455	Sequence 16455, A
38	85	6.0	329	US-11-087-099-8594	Sequence 8594, Ap
39	85	6.0	356	US-11-096-568A-16454	Sequence 16454, A
40	85	6.0	376	US-11-096-568A-16453	Sequence 16453, A
41	84.5	6.0	3655	US-11-075-185-5	Sequence 5, Appl
42	84	6.0	954	US-11-087-099-10843	Sequence 10843, A
43	83.5	5.9	137	US-11-087-099-507	Sequence 507, Ap
44	83.5	5.9	137	US-11-087-099-508	Sequence 508, Ap
45	83.5	5.9	247	US-11-096-568A-12417	Sequence 12417, A

ALIGNMENTS

```
RESULT 1
US-11-087-099-3468
; Sequence 3468, Application US/11087099
; Publication No. US2006004961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087, 099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 3468
; LENGTH: 305
; TYPE: PRT
; ORGANISM: Thermobifida fusca
US-11-087-099-3468

Query Match      14.8%; Score 209; DB 7; Length 305;
Best Local Similarity 28.0%; Pred. No. 8.6e-12;
Matches 84; Conservative 43; Mismatches 133; Indels 40; Gaps 10;

QY 7 PPSIVGIDGSKPVAQALMAVDEAASRDIPRLIYA-----IEPDDREGYAHGAAR 59
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 8 PGTIVVGVDSEGLHLDWAIDAAAGRATRLIVYMGLEPLVTPGGRTPRTPSEVS 67
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 60 KLAAG--ENAVRYAFVAVEAADRPVKVEITQERPTSLIRASAAALVCVGAIGHNF 117
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 68 QAAALLLEALR-----RVQEAAPSLRAVTEVSRAEHALIKSKODDELVSGRSGV 123
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 118 RPERVSTAAALALSAQCPVAIVRPHRPVPIGRDAWITVEADSSDIGVLGAVMAA-R 176
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 124 ASLELGVAAQVAVSHATCPVVPVPTQBAARAGRVRVGVGDSGEHAALRALFAVBAER 183
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 177 LR-----DSVRYVVTGQSGVGTGDDVRAISLDRMLAR--QRRYP-----DV 217
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 184 LRRELVAIVVMAQAPDAVDEPFTVLQADVAVDREQVARRBEMWLRITVDEARTPTQRPV 243
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 218 RVQSAVAH--GELIDYLAGRSVHMVVLASDQ--EHVEQLVGAPNAVLOEAGCTLVV 274
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 244 RVETPEBHPAALALDE---GADADLVVSGRGRGFTGLILSGVQVLAHNAVVAVV 299
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

RESULT 2
US-11-055-822-14
; Sequence 14, Application US/11055822
; Publication No. US20050260707A1
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
```



```

QY 52 -----AAHGAARKLAAMENVRVAFVTEAADRPVKVEITQCRPTSLIRASAAA 105
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 90 YDVTMDASGRIWAPAPAP-----AADRL-----PVVYFHGGGA- 126

QY 106 LVCVGAIGVHRRPERVGSSTAALALSAOCPAVAIVPHRVPIGRDAAMIVTEADGSSDIG 165
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 127 -LFSPLAIFGPNVCRCRLCAALGAVVSVNYRLA--PEHNMWPAID-----DGVDAIR 175

QY 166 VLLGAVMAEAR-----LRDS-FVRVYTCQSGVGDGTDDVRSLSLDRMLARKMPRIYDVAV 219
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 176 FL-----DARGGVPLDGDGVFVDLGTCTPLAGESAGNIVHNVANRMAAAMPASARALRV 229

QY 220 QSAAVHGLDYLVLGGRS 238
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 230 --AGVF-PVQYFGGVERT 245

RESULT 9
US-11-096-568A-13340
; Sequence 13340, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 13340
; LENGTH: 316
; TYPE: PRT
; ORGANISM: Triticum aestivum
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(316)
; OTHER INFORMATION: Ceres Seq. ID no. 15176086
US-11-096-568A-13340

Query Match 6.8%; Score 95.5; DB 7; Length 316;
Best Local Similarity 25.1%; Pred. No. 0.26;
Matches 69; Conservative 29; Mismatches 102; Indels 75; Gaps 11;

QY 28 VDEAASROIPLRL-----LYAIEPDDPGYAAHGAARKLAAMENVRVAFVTEAADRPV 82
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 82 VDLAACRRRGINVTNAGAAFAVDYTDYGVGLVVAIVLRALAAEAHVRAGRMAAD-GDYPL 140

QY 83 KVEVEITQCRPTSLIRASAAALVCVGAIGVHRRPERVGSSTAALALSAOCPAVAIVP 142
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 141 TTKV-----SEKRVGIYGLNIGSRIR--RLAIFGCAVSNSRSPKRSAPY 185

QY 143 HRVPIGRDAAMIVTEADGSSDIGVLLGAVMAEAR-----LR 178
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 186 EFVPAVRDIA-----AGSDVLVLSCLTETTRVHVRNEVEALGKAGVLNVGRGGIV 238

QY 179 DSPRVVYTCQSGV-GDTRDDVRSLSLDRMLARKMPRIYDVRSVOSAAPHGLDYLGLGR 237
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 239 DEP-ELVACLQGVGAGADVPES-----EPVPELPSMDNVVL----- 278

QY 238 SVHNVYLSASDQEHVEQLVGAPGNVLOBAGCTLL 272
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 279 SAHRAVATPESIRGTIELVAANLDAVF--AGKPLL 311

RESULT 10
US-11-096-568A-18547
; Sequence 18547, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides

```

```

; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 18547
; LENGTH: 345
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: m1sc_feature
; LOCATION: (1)..(345)
; OTHER INFORMATION: Ceres Seq. ID no. 12365936
US-11-096-568A-18547

Query Match
Beet Local Similarity 21.9%; Score 94.5; DB 7; Length 345;
Matches 77; Conservative 28; Mismatches 112; Indels 135; Gaps 14,

QY 10 IVGIDGSKPAVOALMAVDEASRDIPLRLLVAIEPDDG--VAAHGAARKLAAENA 67
   : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 1 MAALDLSGSEELMRLEPFIRDASHGSSPLPHHPQLSPSPSPFHQVA-VAAASSYG 59
   : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
QY 68 VRYAFTVAEADRPVKVEVETIOER-----PTSLIRASAA----- 104
   : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 60 GNYPPAARDEAGQLSPSQMOYIQARLHQRRQAQTSVLGPPAQPMSASAPARAPQK 119
   : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
QY 105 ----ALVCGAIG-----VHFR----- 118
   : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 120 LYPRRAAALGVGGGDPAAPAPHPALRHLRHRGGAGGLRPGRLPPRRRGAQLPRLQ 179
   : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
QY 119 -----PR-FVGSFAALLASQCPVAIVRPHR----- 144
   : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 180 RRLPRAPRLRRRQAPPEVPPEHRHVQEGQASLRSRCV-VVRHQOQLLLAVRRRLD 238
   : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
QY 145 -----VPIGDAAMIVVEADGSSD--IGVLLGAVMAEALRDSPPRVVTCRQSGV 192
   : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 239 VLPRVRHVLPVPVPHVAVRLRARDGADAAAGIQRGAVGGRLLRHQVAVVADR----- 294
   : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
QY 193 GDTGDDVVASI--DRWLARWQPRYPDVAVGSA-----VHGE-----LDYLAG 234
   : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 295 --LGIHPRQLIVAPCWLAGMLPSYPHCRBAAGGILKVNGBRRADGIDILIRG 344
   : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :

RESULT 11
US-11-087-099-2546
; Sequence 2546, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 2546
; LENGTH: 161
; TYPE: PRT
; ORGANISM: Glycine max
US-11-087-099-2546

Query Match
Beet Local Similarity 26.7%; Score 93.5; DB 7; Length 161;
Matches 43; Conservative 25; Mismatches 70; Indels 23; Gaps 6,

QY 2 NQSHKPEIIVGIDGSKPAVOALMAVDEASRDIPRLVAIEP---DDPGVAHGA-- 57
   : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 5 SSSSEKQVVLIGIDSEOSTYALMWALDHPPSPFIKVLILHSPRTASAVGFGPGAAE 64
   : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
QY 58 -----ARKLAA--AENAVRYAFTVAEADRPVK--VEVETIOERPTSLIR--SAA 104
   : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 65 ILPIVDSLRKIARVLETAQQLCF-----NKSVDNTAIVVGGDPRLNVCGADVTKYRA 118
   : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
QY 105 ALVCGAIGVHFRPERVGSFAALLASQCPVAIVRPHRV 145

```

Db 119 AMLVVGSHGGAIKRAVLGVSVDYCAHHAHCTVMIVKKPKI 159

RESULT 12

US-11-087-099-6989
 ; Sequence 6989, Application US/11087099
 ; Publication No. US20060041961A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Abad, Mark S. et al.
 ; TITLE OF INVENTION: Genes and Uses for Plant Improvement
 ; FILE REFERENCE: 38-21(53450)B EP
 ; CURRENT APPLICATION NUMBER: US/11/087,099
 ; CURRENT FILING DATE: 2005-03-22
 ; NUMBER OF SEQ ID NOS: 12464
 ; SEQ ID NO 6989
 ; LENGTH: 161
 ; TYPE: PRT
 ; ORGANISM: Glycine max
 US-11-087-099-6989

Query Match 6.6%; Score 93.5; DB 7; Length 161;
 Best Local Similarity 26.7%; Pred. No. 0.16;
 Matches 43; Conservative 25; Mismatches 70; Indels 23; Gaps 6;

QY 2 NQSHKPSIVVIGIDGSKPAVOALMAVDEAASRDIPRLLYAIEP---DDPGYAANGAA- 57

Db 5 SSSSEKQVVLIGIDSEOSTALNMALDHFPSPFKVLHRSRTATSAVGFAGPGLAE 64

QY 58 -----ARKLAA--AENAVRYAFTAVEADRPVK--VEEITDERPVTSLIRA--SAAA 104

Db 65 ILPIVSDLRKTAARVLETAKQLCF-----NKSVDVTAEVDEDPNNVLCDAYDKYRA 118

QY 105 ALVCGAIGVNHFRPERVGTAAALALSAQCPVAIVRHRV 145

Db 119 AILVVGSHGGAIKRAVLGVSVDYCAHHAHCTVMIVKKPKI 159

RESULT 13

US-11-096-568A-18548
 ; Sequence 18548, Application US/11096568A
 ; Publication No. US20060048240A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Alexandrov, Nickolai et al.
 ; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
 ; FILE REFERENCE: 2750-1592PUS2
 ; CURRENT APPLICATION NUMBER: US/11/096,568A
 ; CURRENT FILING DATE: 2005-04-01
 ; NUMBER OF SEQ ID NOS: 34471
 ; SEQ ID NO 18548
 ; LENGTH: 333
 ; TYPE: PRT
 ; ORGANISM: Zea mays subsp. mays
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (1)_(333)
 ; OTHER INFORMATION: Cereb Seq. ID no. 12365697
 US-11-096-568A-18548

Query Match 6.6%; Score 93.5; DB 7; Length 333;
 Best Local Similarity 21.9%; Pred. No. 0.42;
 Matches 73; Conservative 27; Mismatches 96; Indels 137; Gaps 15;

QY 29 DEASRDIPRLLYAIEPDPG--YAAHGAARKLAAMENAVRYAFTAVEADRPVKYEV 86

Db 10 DASAHGSSP--LLPHHQPLSPSPSPFFQAVA--VAAASYGANYFPAAPDEAGQLSPSQM 66

QY 87 EITGR-----PVTSLIRASAAA-----ALVCGAIG----- 113

Db 67 QYIARLHLQRRQACTSVLGFPAQPMKASASAAAPAPAPQKLYPPRAAALAGVGGGDPA 126

QY 114 -----VHHFR----- 118

Db 127 PAQPHPALRLRRRGGGAGLRPGRLPPRRRGAPQLPRQRLLPPRAPRLRRRQAREPV 186

QY 119 PE--RVGSTAALALSAQCPVAIVRHR-----VPIGRDAAM 153

Db 187 PEHRRVQEGROEALRSRSCR--VVRPHQQLLLAVERRRDLVLRVRRHVLVPVHVAV 245

QY 154 IVEADGSSD--IGVLGAVMAEARLDSPVYVTCRQSGVGDGDDVRASL--DRWLAR 209

Db 246 RLARADGADPAALQVAGRGRLRAHQVPVYRDR-----IGLPPRLQIVAPCWLAG 299

QY 210 WQPRYPDVVRVQSA-----VHGE-----LLDYIAG 234

Db 300 WLPSYPHCRRAAGGGLKKVGNRRADGILDILNG 332

RESULT 14

US-11-096-568A-27471
 ; Sequence 27471, Application US/11096568A
 ; Publication No. US20060048240A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Alexandrov, Nickolai et al.
 ; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
 ; FILE REFERENCE: 2750-1592PUS2
 ; CURRENT APPLICATION NUMBER: US/11/096,568A
 ; CURRENT FILING DATE: 2005-04-01
 ; NUMBER OF SEQ ID NOS: 34471
 ; SEQ ID NO 27471
 ; LENGTH: 174
 ; TYPE: PRT
 ; ORGANISM: Arabidopsis thaliana
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (1)_(174)
 ; OTHER INFORMATION: Cereb Seq. ID no. 1817050
 US-11-096-568A-27471

Query Match 6.5%; Score 91; DB 7; Length 174;
 Best Local Similarity 23.1%; Pred. No. 0.3;
 Matches 37; Conservative 30; Mismatches 65; Indels 28; Gaps 6;

QY 10 IVVGIDGSKPAVOALMAVD---EASRDIPRLLYAIEPDPG--- 55

Db 10 VVAVDSDSEVSMELRALNMLKSSSSDSFVLLH--VQPSVAVAGVSGTIPFGSP 68

QY 56 ---AARKLAAMENAVRYAFTAVEADR-----PVYVEITDERPVTSLIRA--SA 102

Db 69 GLEVPFAFTAAIEHQKRTITDITIEHASQICAEKSVSRVNVKTQVIGDPKXKICEAVENTL 128

QY 103 AALVCGAIGVNHFRPERVGTAAALALSAQCPVAIVR 142

Db 129 HADLLVVGSRAYGRKIMPLGVSNSVCTNHNAHCYVLIK 168

RESULT 15

US-10-454-437-254
 ; Sequence 254, Application US/10454437
 ; Publication No. US20050277115A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Pompeius, Markus
 ; APPLICANT: Krogger, Burkhard
 ; APPLICANT: Schroder, Hartwig
 ; APPLICANT: Zeidler, Oskar
 ; APPLICANT: Haberhauser, Gregor
 ; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
 ; FILE REFERENCE: BGI-128CPCN
 ; INVOLVED IN HOMEOSTASIS AND ADAPTATION
 ; CURRENT APPLICATION NUMBER: US/10/454,437
 ; CURRENT FILING DATE: 2003-06-13
 ; PRIOR APPLICATION NUMBER: US 60/141031
 ; PRIOR FILING DATE: 1999-06-25
 ; PRIOR APPLICATION NUMBER: DE 19931636.8

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocelelation Ltd.

OM protein - protein search, using sw model

Run on: March 23, 2006, 04:56:29 ; Search time 90.1288 Seconds
(without alignments)
1652.628 Million cell updates/sec

Title: US-10-617-038-22

Perfect score: 1720
Sequence: 1 MTEPAAWDEGKPRITLTLMN.....TEVGQDQYVHPIVNPASP 339

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq_21: *
1: geneseqp1980s: *
2: geneseqp1990s: *
3: geneseqp2000s: *
4: geneseqp2001s: *
5: geneseqp2002s: *
6: geneseqp2003s: *
7: geneseqp2003bs: *
8: geneseqp2004s: *
9: geneseqp2005s: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1720	100.0	339	6	ABU34644
2	1720	100.0	339	6	ABU36709
3	1720	100.0	339	8	AD137301
4	726	42.2	310	8	ADS30122
5	726	42.2	315	9	ABM95087
6	617.5	35.9	308	8	ADS21176
7	597.5	34.7	505	6	ABU22654
8	587.5	34.2	338	6	ABU20143
9	534.5	31.1	310	4	AAU38983
10	531.5	30.9	309	4	AAU34587
11	531.5	30.9	309	8	ABU28646
12	531.5	30.9	309	8	ADN18198
13	531.5	30.9	311	6	ABU28039
14	530.5	30.3	314	7	ABO63860
15	322	18.7	318	7	ADH88496
16	311	18.1	303	6	ABU17706
17	304	17.7	309	6	ABU24165
18	302	17.6	1249	4	ABG21666
19	299.5	17.4	340	4	AAU49793
20	299.5	17.4	340	6	ABM46312
21	298	17.3	321	7	ADH85822
22	294	17.1	304	6	ADB06322
23	290	16.9	305	4	AAU35194
24	289	16.8	305	6	ABU29145

25	284	16.5	306	6	ABU42743	ABU42743	Protein e
26	284	16.5	312	5	ABP40738	ABP40738	Staphyloc
27	284	16.5	312	8	ADS08044	ADS08044	Staphyloc
28	283	16.5	306	6	ADA89524	ADA89524	Staphyloc
29	283	16.5	306	6	ABU16314	ABU16314	Protein e
30	282	16.5	306	6	ABM72329	ABM72329	Staphyloc
31	282	16.4	303	4	AAU33897	AAU33897	Staphyloc
32	282	16.4	306	4	AAU36707	AAU36707	Staphyloc
33	279.5	16.2	303	9	ADW26675	ADW26675	Fructo-ol
34	279.5	16.2	312	6	ADB07554	ADB07554	Alloisococ
35	279	16.2	330	4	AAH86203	AAH86203	C. glutam
36	279	16.2	330	4	AAH79271	AAH79271	Corynebact
37	279	16.2	330	4	AAH91870	AAH91870	C. glutami
38	279	16.2	330	6	ABP60402	ABP60402	Corynebact
39	279	16.2	330	7	ADD13373	ADD13373	C. glutam
40	278.5	16.2	310	5	ABP30506	ABP30506	Streptococ
41	278.5	16.2	310	8	ADV88590	ADV88590	Streptococ
42	278.5	16.2	310	8	ADV79843	ADV79843	Streptococ
43	278.5	16.2	310	8	ADV81996	ADV81996	Streptococ
44	278.5	16.2	320	5	ABP28871	ABP28871	Streptococ
45	276.5	16.1	302	4	AAU33459	AAU33459	Enterococ

ALIGNMENTS

RESULT 1	
ABU34644	ABU34644 standard; protein; 339 AA.
ID	ABU34644;
AC	ABU34644;
DT	19-JUN-2003 (first entry)
XX	
XX	Protein encoded by Prokaryotic essential gene #20171.
XX	
XX	Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX	
XX	Mycobacterium bovis.
XX	
XX	WO20027183-A2.
XX	
XX	03-OCT-2002.
XX	
XX	21-MAR-2002; 2002WO-US009107.
XX	
XX	21-MAR-2001; 2001US-00815242.
XX	
XX	06-SEP-2001; 2001US-00948993.
XX	
XX	25-OCT-2001; 2001US-0342923P.
XX	
XX	08-FEB-2002; 2002US-00072851.
XX	
XX	06-MAR-2002; 2002US-0362699P.
XX	
XX	(EUIT-) ELITRA PHARM INC.
XX	
XX	Wang J, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
XX	Wall D, Trawick JD, Carr GJ, Yamamoto R, Foreyth RA, Xu HH;
XX	WPI; 2003-029926/02.
XX	N-PSDB; ACA38514.
XX	
XX	New antisense nucleic acids, useful for identifying proteins or screening
XX	for homologous nucleic acids required for cellular proliferation to
XX	isolate candidate molecules for rational drug discovery programs.
XX	
XX	Claim 25; SEQ ID NO 62568; 1766pp; English.
XX	
XX	The invention relates to an isolated nucleic acid comprising any one of
XX	the 6213 antisense sequences given in the specification where expression
XX	of the nucleic acid inhibits proliferation of a cell. Also included are:
XX	(1) a vector comprising a promoter operably linked to the nucleic acid
XX	encoding a polypeptide whose expression is inhibited by the antisense
XX	nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX	polypeptide or its fragment whose expression is inhibited by the

Query	Match	100.0%;	Score 1720;	DB 6;	Length 339;				
Best Local Similarity	100.0%;	Pred. No. 3,1e-167;							
Matches 339;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;					
QY	1	MTEPAMDEGKPRITLTMPALDITTSVDVVRPTEKMRCGAPRYDGGGGINVARIYHV	60						
DB	1	MTEPAMDEGKPRITLTMPALDITTSVDVVRPTEKMRCGAPRYDGGGGINVARIYHV	60						
QY	61	LGCGSTALPPAGGSGSLMLLGNAGVPRFRTPLAASSTRESFTYNEBRTAKQYRFVLPG	120						
DB	61	LGCGSTALPPAGGSGSLMLLGNAGVPRFRTPLAASSTRESFTYNEBRTAKQYRFVLPG	120						
QY	121	PSLTVAEBOECIDETRGAAASAFVVASGSLPPGAAAYQYORVADICRRSSTPLILDTSG	180						
DB	121	PSLTVAEBOECIDETRGAAASAFVVASGSLPPGAAAYQYORVADICRRSSTPLILDTSG	180						
QY	181	GGLQHIISSGVFLKASVRELRCVSESLITTEBEOQLAAAHELIDRGRAVWVVSIGSGQAL	240						
DB	181	GGLQHIISSGVFLKASVRELRCVSESLITTEBEOQLAAAHELIDRGRAVWVVSIGSGQAL	240						
QY	241	LATRHASHRFPSSIPMTAVSGCAGDAPYAAITVGSIRGSKLTKSVRLGNAAGAMLTPG	300						
DB	241	LATRHASHRFPSSIPMTAVSGCAGDAPYAAITVGSIRGSKLTKSVRLGNAAGAMLTPG	300						
QY	301	TAACNRDDVERFFELAAEPTEVGQDYQYWHPIVNEASP	339						
DB	301	TAACNRDDVERFFELAAEPTEVGQDYQYWHPIVNEASP	339						
RESULT 2									
ABU36709									
ID	ABU36709	standard: protein; 339 AA.							
AC	ABU36709;								
XX	19-JUN-2003	(first entry)							
XX	DT								
DE	Protein encoded by Prokaryotic essential gene #22236.								
XX									
KM	Antisense; prokaryotic essential gene; cell proliferation; drug design.								
XX									
OS	Mycobacterium tuberculosis.								
XX									
FN	WO200277183-A2.								
XX									
PD	03-OCT-2002.								

[illegible]

QY 181 GGLQHISSGVFLAKASVRELRCEVSELTTEPEQLAAAHLEIDRGRAEVVVVSLGSGCAL 240
 DB 181 GGLQHISSGVFLAKASVRELRCEVSELTTEPEQLAAAHLEIDRGRAEVVVVSLGSGCAL 240
 QY 241 LATTHASHRFSISPMITAVSGVAGADAMVAATVGLSRGWSLTKSVRLGNAGAAMLLTPG 300
 DB 241 LATTHASHRFSISPMITAVSGVAGADAMVAATVGLSRGWSLTKSVRLGNAGAAMLLTPG 300
 QY 301 TAACNRDVERFFELAAPEFTEVGODQYVMHPINVPEASP 339
 DB 301 TAACNRDVERFFELAAPEFTEVGODQYVMHPINVPEASP 339

RESULT 3
 ADI37301
 ID ADI37301 standard; protein; 339 AA.

AC ADI37301;
 XX
 DT 22-APR-2004 (first entry)

DE M. tuberculosis low oxygen induced antigen Rv2029c SEQ ID NO:22.

XX
 KM mycobacterial infection; vaccine; tuberculosis;
 KM mycobacterium tuberculosis; immunisation; antibacterial; gene therapy;
 KM low oxygen induced antigen.
 OS Mycobacterium tuberculosis.

XX W0204006952-A2.

XX 22-JAN-2004.

XX 08-JUL-2003; 2003WO-DK000477.

XX 13-JUL-2002; 2002DK-00001098.

XX (STAT-) STATENS SERUM INST.

XX Andersen P, Rosenkrands I, Stryhn A;

XX WPI; 2004-122778/12.

XX N-PSDB; ADI37346.

PT Use of one or more polypeptides or their fragments, which are expressed during the latent stage of the mycobacterial infection, and/or nucleic acids encoding the polypeptides, for a therapeutic vaccine against tuberculosis.
 PT
 PT

PS Claim 3; SEQ ID NO 22; 76pp; English.

XX The present invention describes polypeptides or their fragments, which are expressed during the latent stage of a mycobacterial infection, and/or nucleic acids encoding the polypeptides, which are useful for creating a therapeutic vaccine against tuberculosis. Also described: (1) a therapeutic vaccine against tuberculosis comprising one or more polypeptides; (2) a method for treating an animal, including a human being, with tuberculosis caused by virulent mycobacteria, e.g. by Mycobacterium tuberculosis, M. africanum or M. bovis; (3) a method for immunising an animal, including a human being, against tuberculosis caused by virulent mycobacteria; (4) a method of diagnosing tuberculosis caused by virulent mycobacteria in an animal, including a human being; (5) a method for diagnosing previous or ongoing infection with a virulent mycobacterium; and (6) a method of diagnosing Mycobacterium tuberculosis infection in a subject. The polypeptides have antibacterial activities, and can be used in vaccines and in gene therapy. The polypeptides are useful for the manufacture of a therapeutic vaccine for treating an individual who is infected by a virulent mycobacterium, e.g. M. tuberculosis, and who is not vaccinated with BCG against tuberculosis. The present sequence represents a low oxygen induced antigen, which is used in the exemplification of the present invention.

XX Sequence 339 AA;

Query Match 100.0%; Score 1720; DB 8; Length 339;
 Best Local Similarity 100.0%; Pred. No. 3.1e-167;
 Matches 339; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTEPAAWDEKPRITLTVMNPALDITTSVDVVRPTKMRGAPRYDGGGINVARIYHV 60
 DB 1 MTEPAAWDEKPRITLTVMNPALDITTSVDVVRPTKMRGAPRYDGGGINVARIYHV 60
 QY 61 LGGCTALFPAGSGTSLMALIGDAGVPPRVPIAASTRSFVNESRTAKOYRFVLP 120
 DB 61 LGGCTALFPAGSGTSLMALIGDAGVPPRVPIAASTRSFVNESRTAKOYRFVLP 120
 QY 121 PSLTVAEQECCLDELRGAAASAAFFVVASGSLPEGVADYYQVADICRSSTPLIDTSG 180
 DB 121 PSLTVAEQECCLDELRGAAASAAFFVVASGSLPEGVADYYQVADICRSSTPLIDTSG 180
 QY 181 GGLQHISSGVFLAKASVRELRCEVSELTTEPEQLAAAHLEIDRGRAEVVVVSLGSGCAL 240
 DB 181 GGLQHISSGVFLAKASVRELRCEVSELTTEPEQLAAAHLEIDRGRAEVVVVSLGSGCAL 240
 QY 241 LATTHASHRFSISPMITAVSGVAGADAMVAATVGLSRGWSLTKSVRLGNAGAAMLLTPG 300
 DB 241 LATTHASHRFSISPMITAVSGVAGADAMVAATVGLSRGWSLTKSVRLGNAGAAMLLTPG 300
 QY 301 TAACNRDVERFFELAAPEFTEVGODQYVMHPINVPEASP 339
 DB 301 TAACNRDVERFFELAAPEFTEVGODQYVMHPINVPEASP 339

RESULT 4
 ADS30122
 ID ADS30122 standard; protein; 310 AA.

XX ADS30122;

XX 02-DEC-2004 (first entry)

XX Bacterial polypeptide #19155.

XX Recombinant DNA construct; transformed plant; improved plant property;
 KM cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
 KM pathogen tolerance; pest tolerance; plant disease resistance;
 KM cell cycle pathway modification; plant growth regulator;
 KM homologous recombination; seed oil yield; protein yield; carbohydrate;
 KM nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
 KM bacterial polypeptide.

XX Bacteria.

XX US2003233675-A1.

XX 18-DEC-2003.

XX 20-FEB-2003; 2003US-00369493.

XX 21-FEB-2002; 2002US-0360039P.

XX (CAOY/) CAO Y.
 XX (HINK/) HINKLE G J.
 XX (SLAT/) SLATER S C.
 XX (CHEN/) CHEN X.
 XX (GOLD/) GOLDMAN B S.

XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;

XX WPI; 2004-061375/06.

PT New recombinant DNA construct comprising a promoter positioned to provide for expression of a polynucleotide encoding a polypeptide from a microbial source, useful for producing plants with improved properties.

PS Claim 1; SEQ ID NO 19155; 122pp; English.

KW pathogen tolerance; pest tolerance; plant disease resistance;
 KW cell cycle pathway modification; plant growth regulator;
 KW homologous recombination; seed oil yield; protein yield; carbohydrate;
 KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
 KW bacterial polypeptide.
 XX Bacteria.
 OS
 PN US2003233675-A1.
 PD 18-DEC-2003.
 XX
 XX 20-FEB-2003; 2003US-00369493.
 XX
 XX 21-FEB-2002; 2002US-0360039P.
 XX
 PA (CAOY/) CAO Y.
 PA (HINK/) HINKLE G J.
 PA (SLAT/) SLATER S C.
 PA (CHEN/) CHEN X.
 PA (GOLD/) GOLDMAN B S.
 XX
 PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
 XX
 DR WPI; 2004-061375/06.
 XX
 PT New recombinant DNA construct comprising a promoter positioned to provide
 PT for expression of a polynucleotide encoding a polypeptide from a
 PT microbial source, useful for producing plants with improved properties.
 XX
 XX Claim 1; SEQ ID NO 10209; 122pp; English.
 XX
 CC The invention relates to a recombinant DNA construct comprising a
 CC promoter functional in a plant cell, where the promoter is positioned to
 CC provide for expression of a polynucleotide encoding a polypeptide from a
 CC microbial source. The invention also relates to a transformed plant
 CC comprising the recombinant DNA construct and a method of producing a
 CC transformed plant having an improved property. The plant is a crop plant
 CC such as maize or soybean. The method of producing a transformed plant
 CC having an improved property comprises transforming a plant with the
 CC recombinant DNA construct and growing the transformed plant, where the
 CC polynucleotide or polypeptide is useful for improving plant properties.
 CC The recombinant DNA construct is useful for producing plants with
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
 CC increased resistance to plant disease, better growth rate by modification
 CC of the cell cycle pathway with plant growth regulators, increased rate of
 CC homologous recombination, modified seed oil or protein yield and/or
 CC content, improved yield by modification of carbohydrate, nitrogen or
 CC phosphorus use and/or uptake, by modification of photosynthesis or by
 CC providing improved plant growth and development under at least one stress
 CC condition, improved lignin production or improved galactomannan
 CC production. This sequence represents a bacterial polypeptide used in the
 CC scope of the invention. Note: The sequence data for this patent did not
 CC form part of the printed specification but was obtained in electronic
 CC format from USPTO at seqdata.uspto.gov/sequence.html.
 XX
 SQ Sequence 308 AA:
 Query Match 35.9%; Score 617.5; DB 8; Length 308;
 Best Local Similarity 43.0%; Pred. No. 3.3e-54;
 Matches 131; Conservative 55; Mismatches 118; Indels 1; Gaps 1;
 QY 11 KPRITITLMPALDITTSVDVVRPREKRCAPYDPCGGGGINVARIYHVGCGSTLFP 70
 DB 2 KSSIFITLTLSPAFKDSIVSVESVQPHKRCSPYQPGGGINVRAIKKLNSSVALYP 61
 QY 71 AGSGTSLTLMALGDAVFPFVPIASTRESFTVNESRTAKOYFVLPGSLTVAEQEQ 130
 DB 62 ICGTGTHFRKLLTBEGIEQITTPQIKMTRENFIIVESTNNQRFGRGSEIYKEEHEQ 121
 QY 131 CLDELRGAAASAAFPVAVSGSLPPGVADYQYRVADICRSTPLILDTSGGGLQH-ISSG 189

DB 122 MIAVIGDPSNDFEYFVVGSGTPKGVDPDEFYTRLSQIILQKARLILDTSGAALKHCKEG 181
 QY 190 VFLLKASVRELRBCVGSSELTTEPEQLAAAHELLDRGAEEVNVVSLGQALLATRNAHSR 249
 DB 182 IFLLCKPWINSELSELVIGELITTKGQEEAAMEIINSKVELILVYSLGAGAFVLSKQGIYH 241
 QY 250 FSSITPTAVSGVAGDAMVAIIVGLSRGWSLKSIVRLGNVAGAAAMLLTGTACNRDDV 309
 DB 242 VTAPPVKRSSTVGAGDSMVAGMYLSRGNHYEDVYIKFVAGACGTAAATMNFTELCXEDV 301
 QY 310 ERFFE 314
 DB 302 ERLYK 306
 RESULT 7
 ABU22654
 ID ABU22654 standard; protein; 505 AA.
 XX
 AC ABU22654;
 XX
 DT 19-JUN-2003 (first entry)
 XX
 DE Protein encoded by Prokaryotic essential gene #8181.
 XX
 KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
 XX
 OS Burkholderia mallei.
 XX
 FN W0200277183-A2.
 PD 03-OCT-2002.
 XX
 XX 21-MAR-2002; 2002WO-US009107.
 XX
 XX 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX
 PI Wang L, Zamudio C, Malone C, Haeslbeck R, Ohlsen KL, Zyskind JW,
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Foreyth RA, Xu HH;
 XX
 DR WPI; 2003-022926/02.
 DR N-PSDB; ACA26524.
 PT
 PT New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX
 XX Claim 25; SEQ ID NO 50578; 1766pp; English.
 XX
 CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-regulated gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene

CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 505 AA;

Query Match 34.7%; Score 597.5; DB 6; Length 505;
Best Local Similarity 41.9%; Pred. No. 7,66-52;
Matches 127; Conservative 59; Mismatches 116; Indels 1; Gaps 1;

QY 12 PRITLTNNPALDITTSVDVVRPTEKMRGAPRYDPGGGINVARIYHVLGGCSTALFPA 71
DB 194 PEIVTLTPNPAIDVATCVERVTDRKLCRCGPARDPGGGGINVARIYHVLGGCSTALFPA 253
QY 72 GGSGSLMALGAGCPFRVPIPIAASRSTFNESTAKQYRFLPSPSLTVAEQEOC 131
DB 254 GGGGGLRLGLLDEGVRAMGIDTAGETRENFSVLESTGREFVFLPGPILAAHEWPRC 313
QY 132 LDELRAAASAFVAVASGLPPGVAADYQYRVADICRRSSPTLIDTSGGL-CHISSGV 190
DB 314 VEALGRADASRYVMGSGSLPPGMPDDCYALARASRGRTVYDTSGLPILAAALDAGV 373
QY 191 FLKASVRELRCEVSELTTEPEQLAAAHEDIGRAEVVVVSLGSGCALIATRHASHRF 250
DB 374 YLVPSIGELRALTGLPLEDDGARLAAARAIVAGGAAQIVALTIGDAGALVSRDDAVRL 433
QY 251 SSIMTAVSGVAGADANVAATITVLSRGWSIKSVRLGNAAGAMLTLPGRACNRDVE 310
DB 434 PGVAVARSALIGADDSFVAGLVAAALNRGANVADARRHALLAASASLSTGTGALTKEDIA 493
QY 311 RFF 313
DB 494 RIV 496

RESULT 8

ABU20143
ID ABU20143 standard; protein, 338 AA.

XX ABU20143;

XX 19-JUN-2003 (first entry)

DE Protein encoded by prokaryotic essential gene #5670.

KM Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX *Borrelia cepacia*.

PN WO200277183-A2.

PD 03-OCT-2002.

PF 21-MAR-2002; 2002WO-US009107.

PR 21-MAR-2001; 2001US-00815242.

PR 06-SEP-2001; 2001US-00948993.

PR 25-OCT-2001; 2001US-0342923P.

PR 08-FEB-2002; 2002US-00072851.

PR 06-MAR-2002; 2002US-0362699P.

XX (ELIT-) ELITRA PHARM INC.

PI Wang L, Zamudio C, Malone C, Haebeck R, Ohlsen KU, Zyskind JW;
PI Wall D, Trawick UD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX WPI; 2003-029926/02.
DR N-PSDB; ACA24013.
XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.

XX Claim 25; SEQ ID NO 48067; 1766pp; English.

CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 338 AA;

Query Match 34.2%; Score 587.5; DB 6; Length 338;
Best Local Similarity 43.0%; Pred. No. 4,5e-51;
Matches 132; Conservative 48; Mismatches 124; Indels 3; Gaps 3;

QY 14 IITLTNNPALDITTSVDVVRPTEKMRGAPRYDPGGGINVARIYHVLGGCSTALFPAAG 73
DB 31 IYVTLNPAVDVATVSEHIVDTHTKLCARPRRDPGGGINVARTVHRLGGDVALYIAGG 90
QY 74 STGSLMALGAGCPFRVPIPIAASRSTFNESTAKQYRFLPSPSLTVAEQEOC 133
DB 91 PTGDVLTALAEHLBBERIRIAGETRENVCVETATNGRFLMPGLPGLTGAEMRACAA 150
QY 134 ELRGAASAFVAVASGLPPGVAADYQYRVADICRRSSPTLIDTSGGGLQ-HISSGVL 192
DB 151 RTDAMPSPRYVLVSLGSLPPGAPDDLYATLARAKARGSRVVDDAAGRALQALKKGVHL 210
QY 193 LKASVRELRCEVSELTTEPEQLAAAHEDIGRAEVVVVSLGSGCALIATRHASHRFSS 252
DB 211 VKPSIGELSLAG-EPIDTSACRKASELVARGQADIVALTIGARGGVVTVTRDRLPLG 269
QY 253 IYPTAVSGVAGADANVAATITVLSRGWSIKSVRLGNAAGAMLTLPGRACNRDVERP 312
DB 270 RPAVCSYVAGADDSFVAGVWVWALACGVPDDACRYALAAASAVSRPTALCTRDVERI 329
QY 313 F-ELAAE 318
DB 330 HAEIVAG 336

```
RESULT 9
AAU38383
ID  AAU38383 standard; protein; 310 AA.
AC
XX  AAU38383;
XX
XX  14-FEB-2002 (first entry)
XX
XX  Salmomella typhi cellular proliferation protein #274.
DE
XX
XX  Antisense; prokaryotic cellular proliferation protein; antibiotic;
XX  antibacterial; drug design.
XX
XX  Salmomella typhi.
OS
XX  WO200170955-A2.
XX
XX  27-SEP-2001.
XX
XX  21-MAR-2001; 2001WO-US009180.
XX
XX  21-MAR-2000; 2000US-0191078P.
XX  23-MAY-2000; 2000US-0206848P.
XX  26-MAY-2000; 2000US-0207727P.
XX  23-OCT-2000; 2000US-0242578P.
XX  27-NOV-2000; 2000US-0253625P.
XX  22-DEC-2000; 2000US-0257931P.
XX  16-FEB-2001; 2001US-0269308P.
XX
XX  (ELIT-) ELITRA PHARM INC.
XX
XX  Haselbeck R, Ohlsen KU, Zyskind JW, Wall D, Trawick JD, Carr GJ;
XX  Yamamoto RT, Xu HH;
XX  WPI; 2001-611495/70.
XX  N-PSDB; AAS56242.
XX
XX  New polynucleotides for the identification and development of
XX  antibiotics, comprise sequences of antisense nucleic acids.
XX
XX  Example 3; SEQ ID NO 13976; 511pp; English.
XX
XX  The invention relates to antisense inhibitors of genes essential to
XX  prokaryotic cellular proliferation, their use in identifying the genes,
XX  their use in the discovery of novel antibiotics, the essential genes
XX  themselves and the encoded proteins. The prokaryotes used are Escherichia
XX  coli, Staphylococcus aureus, Salmomella typhi, Klebsiella pneumoniae,
XX  Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
XX  useful for the identification of potential new targets for antibiotic
XX  development. The antisense nucleic acids can also be used to identify
XX  proteins used in proliferation, to express these proteins, and to obtain
XX  antibodies capable of binding to the expressed proteins. The proteins can
XX  be used to screen compounds in rational drug discovery programmes. The
XX  antisense nucleic acid sequence is also useful to screen for homologous
XX  nucleic acids which are required for cell proliferation in a wide variety
XX  of organisms. The present sequence represents an essential prokaryotic
XX  cellular proliferation protein. Note: The sequence data for this patent
XX  did not form part of the printed specification, but was obtained in
XX  electronic format directly from WIPO at
XX  ftp.wipo.int/pub/published_pct_sequences
XX
XX  Sequence 310 AA;
XX
XX  Query Match 31.1%; Score 534.5; DB 4; Length 310;
XX  Best Local Similarity 41.4%; Pred. No. 1.1e-45;
XX  Matches 126; Conservative 54; Mismatches 119; Indels 5; Gaps 4;
XX
XX  13 RIITITMPALDITTSVUVAPTEKMGRCAPRYDGGGGINVARVYHVLGGCSTLFPAG 72
XX  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
XX  3 RIVTITLAPSDSATITPQIYPEGKLRCSAPVFEPGGGGINVARIAIHGTAIFPAG 62
```

```
QY 73 GSTGSLMALIGDAGVFPFRVPIPIASTRESPTVNESRTAQYRRVLPGPSITVAEQEQL 132
DB 63 GATGEHLVALLADENVEVSTVDAKMDTRONLHVHVESSGRQRYRVMGATLDDDEFQ-L 121
QY 133 DELGGAASAFAFVVAAGSLPBGVAADYYGRVADICRBSSTFLIDTGGGGIQTSS--GV 190
DB 122 EEQVLEITESGAILVITSSLPBGVVEKLTQILISAQKQGRICITDSSGDALTALALGDI 181
QY 191 FLTKASVRELRRECVCSEILTEPEQL-AAAHELIDRGRAEVVVVSLGSGALLATRHASHR 249
DB 182 ELVKNLKELSALVNRD-LIQPDVVRKAQELVSGKARVYVSLGSGALGISEMCIQ 240
QY 250 FSSITPMTRAVSGVAGDMVAALITVLSRGWSLIRSVPLGNNAAGAAMLITPTAACNRDV 309
DB 241 VVPPPVXSQSTVGAGDSVMGAMTKLAQDASLEEMVRFVGAAGSAATLNOGTRLCSHDDT 300
QY 310 ERFP 313
DB 301 QKIT 304
XX
XX  RESULT 10
XX  AAU34587
XX  ID  AAU34587 standard; protein; 309 AA.
XX  AC
XX  AAU34587;
XX
XX  14-FEB-2002 (first entry)
XX
XX  E. coli cellular proliferation protein #168.
XX
XX  Antisense; prokaryotic cellular proliferation protein; antibiotic;
XX  antibacterial; drug design.
XX
XX  Escherichia coli.
XX
XX  WO200170955-A2.
XX
XX  27-SEP-2001.
XX
XX  21-MAR-2001; 2001WO-US009180.
XX
XX  21-MAR-2000; 2000US-0191078P.
XX  23-MAY-2000; 2000US-0206848P.
XX  26-MAY-2000; 2000US-0207727P.
XX  23-OCT-2000; 2000US-0242578P.
XX  27-NOV-2000; 2000US-0253625P.
XX  22-DEC-2000; 2000US-0257931P.
XX  16-FEB-2001; 2001US-0269308P.
XX
XX  (ELIT-) ELITRA PHARM INC.
XX
XX  Haselbeck R, Ohlsen KU, Zyskind JW, Wall D, Trawick JD, Carr GJ;
XX  Yamamoto RT, Xu HH;
XX  WPI; 2001-611495/70.
XX  N-PSDB; AAS52446.
XX
XX  New polynucleotides for the identification and development of
XX  antibiotics, comprise sequences of antisense nucleic acids.
XX
XX  Example 3; SEQ ID NO 10180; 511pp; English.
XX
XX  The invention relates to antisense inhibitors of genes essential to
XX  prokaryotic cellular proliferation, their use in identifying the genes,
XX  their use in the discovery of novel antibiotics, the essential genes
XX  themselves and the encoded proteins. The prokaryotes used are Escherichia
XX  coli, Staphylococcus aureus, Salmomella typhi, Klebsiella pneumoniae,
XX  Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
XX  useful for the identification of potential new targets for antibiotic
XX  development. The antisense nucleic acids can also be used to identify
XX  proteins used in proliferation, to express these proteins, and to obtain
XX  antibodies capable of binding to the expressed proteins. The proteins can
```


XX ADN18198;
 AC 02-DEC-2004 (first entry)
 DT Bacterial polypeptide #851.
 DE
 XX
 XX Recombinant DNA construct; transformed plant; improved plant property;
 KM cold tolerance; heat tolerance; drought tolerance; herbicide; osmotic;
 KM pathogen tolerance; pest tolerance; plant disease resistance;
 KM cell cycle pathway modification; plant growth regulator;
 KM homologous recombination; seed oil yield; protein yield; carbohydrate;
 KM nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
 KM bacterial polypeptide.
 XX
 OS Bacteria.
 XX
 XX US2003233675-A1.
 XX
 XX 18-DEC-2003.
 XX
 XX 20-FEB-2003; 2003US-00369493.
 XX
 XX 21-FEB-2002; 2002US-0360039P.
 XX
 XX (CAOY/) CAO Y.
 PA (HINK/) HINKLE G J.
 PA (SLAT/) SLATER S C.
 PA (CHEN/) CHEN X.
 PA (GOLD/) GOLDMAN B S.
 XX
 XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
 DR WPI; 2004-061375/06.
 XX
 XX New recombinant DNA construct comprising a promoter positioned to provide
 PT for expression of a polynucleotide encoding a polypeptide from a
 PT microbial source, useful for producing plants with improved properties.
 XX
 PS Claim 1; SEQ ID NO 851; 122pp; English.

CC The invention relates to a recombinant DNA construct comprising a
 CC promoter functional in a plant cell, where the promoter is positioned to
 CC provide for expression of a polynucleotide encoding a polypeptide from a
 CC microbial source. The invention also relates to a transformed plant
 CC comprising the recombinant DNA construct and a method of producing a
 CC transformed plant having an improved property. The plant is a crop plant
 CC such as maize or soybean. The method of producing a transformed plant
 CC having an improved property comprises transforming a plant with the
 CC recombinant DNA construct and growing the transformed plant, where the
 CC polynucleotide or polypeptide is useful for improving plant properties.
 CC The recombinant DNA construct is useful for producing plants with
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
 CC increased resistance to plant disease, better growth rate by modification
 CC of the cell cycle pathway with plant growth regulators, increased rate of
 CC homologous recombination, modified seed oil or protein yield and/or
 CC content, improved yield by modification of carbohydrate, nitrogen or
 CC phosphorus use and/or uptake, by modification of photosynthesis or by
 CC providing improved plant growth and development under at least one stress
 CC condition, improved lignin production or improved galactomannan
 CC production. This sequence represents a bacterial polypeptide used in the
 CC scope of the invention. Note: The sequence data for this patent did not
 CC form part of the printed specification but was obtained in electronic
 CC format from USPTO at seqdata.uspto.gov/sequence.html.
 XX
 XX Sequence 309 AA;
 QY
 Query March 30.9%; Score 531.5; DB 8; Length 309;
 Best Local Similarity 41.1%; Pred. No. 2.2e-45;
 Matches 125; Conservative 55; Mismatches 119; Indels 5; Gaps 4;
 13 RIITLTNMPALDITTSVDVVRPTKMRGAPRYPGGGGINVARIHVLAGCSTALPAPAG 72

DB 3 RIYTLTAPSLDSATITPQIYBEGLKRCAPVFEPPGGGINVAAIAHLGSAVATIPFAG 62
 QY GSTSLMALIGDAGVFPFVPIAASTRESPTVESRTAAQYRFVLPGPSLTVAEBOCL 132
 DB 63 GATSEHLVSLADENVAVATVEACDWTRQNLHNVHVEASGEQYRFVMGALNDEBFQ-L 121
 QY 133 DELGGAASAFAVVASGLPFGVAADYYORVADICRRSSTPLIDTSGGLQHTSS--GV 190
 DB 122 EQGLTIESGALIVISGSLPPGVLEKTLQLSAQAQGRICIVDSGSEALSAALATGNI 181
 QY 191 FLKAASRELRCEGSELTEPEOL-AAAHFLDRGAENVVYVSLSGQALATRRHSHR 249
 DB 182 ELVKNQKELSAVNRB-LTQPDVRRKAAQEIYVNSGAKRVVSLGQALGVDSENCIO 240
 QY 250 FSSIPTAVSGVGADMAVVAITVGLSRGWSLTKSVRLGNAAAMLLTPTACANRDV 309
 DB 241 VVPPFVKSQSTVAGDSMVGAMTKLAEANSLSEMAVAFVAAAGSAATLNGSTRCSHDT 300
 QY 310 ERPF 313
 DB 301 QKIV 304

RESULT 13
 ABU28039
 ID ABU28039 standard; protein; 311 AA.
 XX
 XX ABU28039;
 AC
 XX 19-JUN-2003 (first entry)
 DT
 XX Protein encoded by Prokaryotic essential gene #13566.
 DE
 XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
 KM
 XX Enterobacter cloacae.
 OS
 XX W0200277183-A2.
 XX
 XX 03-OCT-2002.
 PD
 XX 21-MAR-2002; 2002WO-US009107.
 XX
 XX 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX
 XX (ELIT-) ELITRA PHARM INC.
 PA
 XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zykind JW;
 PI Wall D, Trawick JB, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 DR WPI; 2003-029926/02.
 XX
 DR N-PSDB; ACA31909.
 XX
 XX Claim 25; SEQ ID NO 55963; 1766pp; English.

CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: March 23, 2006, 05:11:18 ; Search time 13.4872 Seconds
(without alignments)
2418,401 Million cell updates/sec

Title: US-10-617-038-22
Perfect score: 1720
Sequence: 1 MTEPAAWDEGKPRITLTLMN.....TEVGODQYVWHPIVNPASP 339

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: p1r1:0:1
2: p1r2:0:1
3: p1r3:0:1
4: p1r4:0:1

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1720	100.0	339	2 D70942	probable pfkB prot
2	534.5	31.1	310	2 AB0707	6-phosphofructokin
3	531.5	30.9	309	1 KIECB	6-phosphofructokin
4	530.5	30.8	309	2 A85781	6-phosphofructokin
5	530.5	30.8	309	2 E90932	6-phosphofructokin
6	318.5	18.5	315	2 C75626	1-phosphofructokin
7	309.5	18.0	308	2 B97263	tagatose-6-phospha
8	307	17.8	310	2 E43258	tagatose-6-phospha
9	304	17.7	309	2 B96928	1-phosphofructokin
10	300.5	17.5	303	2 A69627	fructose-1-phospha
11	283	16.5	306	2 D89841	fructose-1-phospha
12	275	16.0	307	2 A11735	fructose-1-phospha
13	268	15.6	318	2 A40944	1-phosphofructokin
14	267.5	15.6	310	2 S04358	1-phosphofructokin
15	266.5	15.5	310	2 D90015	1-phosphofructokin
16	266	15.5	306	2 C83753	fructose-1-phospha
17	266	15.5	307	2 A11366	fructose-1-phospha
18	263.5	15.3	307	2 E70178	fructose-1-phospha
19	257	14.9	309	2 A98006	tagatose-6-phospha
20	255.5	14.9	310	2 C39778	tagatose-6-phospha
21	250	14.5	309	2 C95138	tagatose-6-phospha
22	243.5	14.2	317	2 C82450	1-phosphofructokin
23	242	14.1	303	2 B95101	1-phosphofructokin
24	239	13.9	303	2 C97969	1-phosphofructokin
25	227	13.2	316	1 A39407	1-phosphofructokin
26	226	13.1	304	2 AB0888	probable carbohydr
27	223.5	13.0	312	2 AE0783	1-phosphofructokin
28	222.5	12.9	312	2 F85855	fructose-1-phospha
29	222.5	12.9	312	2 B37245	1-phosphofructokin

30	222.5	12.9	312	2 D91011	fructose-1-phospha
31	222	12.9	314	2 A83200	1-phosphofructokin
32	218.5	12.7	312	2 AF0158	1-phosphofructokin
33	206.5	12.0	305	2 F86744	tagatose-6-phospha
34	198	11.5	313	2 A64069	1-phosphofructokin
35	195	11.3	303	2 D75288	carbohydrate kinase
36	187.5	10.9	319	2 H72327	sugar kinase, pfkB
37	174.5	10.1	310	2 AD1707	1-phosphofructokin
38	173.5	10.1	310	2 AG1336	1-phosphofructokin
39	170	9.9	483	2 F87700	rfae protein (lipo
40	153	8.9	313	2 AG3520	ribokinase (EC 2.7
41	152.5	8.9	299	2 E72311	ribokinase - Therm
42	151	8.8	303	2 B87367	ribokinase (import
43	147	8.5	308	2 A10001	ribokinase (EC 2.7
44	144	8.4	319	2 AC00942	probable carbohydr
45	142	8.3	311	2 E75385	carbohydrate kinase

ALIGNMENTS

RESULT 1

D70942
probable pfkB protein - Mycobacterium tuberculosis (strain H37Rv)

C:Species: Mycobacterium tuberculosis

C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004

C:Accession: D70942

R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.,
; Connor, R.; Davies, R.; Devlin, K.; Feltham, T.; Gentles, S.; Hamlin, N.; Holroyd, S.,
; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skellton, S.; Squares, S.

Nature 393, 537-544, 1998

A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A:Reference number: A70500; PMID:98295987; PMID:9634230

A:Accession: D70942

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-339 <COL>

A:Cross-references: UNIPROT:O86352; UNIPARC:UPI00000D1173; GB:AL021899; GB:AL123456; NID

A:Experimental source: strain H37Rv

C:Genetics:

A:Gene: pfkB

C:Superfamily: 6-phosphofructokinase 2

Query Match	100.0%	Score 1720;	DB 2;	Length 339;
Best Local Similarity	100.0%	Pred. No. 7.7e-117;		
Matches 339;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MTEPAAWDEGKPRITLTLMN	2	PSLTVAEQECCLDELRCGAASAAFAVVASGSLPGVAADYQRYADICRSTPLIDTSG 180
DB	1	MTEPAAWDEGKPRITLTLMN	2	PSLTVAEQECCLDELRCGAASAAFAVVASGSLPGVAADYQRYADICRSTPLIDTSG 180
QY	61	LGCGSTLFPAGSGTSGLLMALLGDA	61	LGCGSTLFPAGSGTSGLLMALLGDA
DB	61	LGCGSTLFPAGSGTSGLLMALLGDA	61	LGCGSTLFPAGSGTSGLLMALLGDA
QY	121	PSLTVAEQECCLDELRCGAASAAFAVVASGSLPGVAADYQRYADICRSTPLIDTSG 180		
DB	121	PSLTVAEQECCLDELRCGAASAAFAVVASGSLPGVAADYQRYADICRSTPLIDTSG 180		
QY	181	GGLQHSISGVFLKASVRELCVGS	181	GGLQHSISGVFLKASVRELCVGS
DB	181	GGLQHSISGVFLKASVRELCVGS	181	GGLQHSISGVFLKASVRELCVGS
QY	241	LATRHASHRFSSTIPMTN	241	LATRHASHRFSSTIPMTN
DB	241	LATRHASHRFSSTIPMTN	241	LATRHASHRFSSTIPMTN
QY	301	TAACNRDVERFPFLAAPE	301	TAACNRDVERFPFLAAPE
DB	301	TAACNRDVERFPFLAAPE	301	TAACNRDVERFPFLAAPE

RESULT 2

B0707
6-phosphofructokinase isozyme [imported] - *Salmonella enterica* subsp. *enterica* serovar Typhimurium
C:Species: *Salmonella enterica* subsp. *enterica* serovar Typhi
A:Note: this species has also been called *Salmonella typhi*
C>Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C:Accession: AB0707
R:Pakhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Mole, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.
Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* serovar Typhimurium
A:Reference number: AB0502; MUID:21534947; PMID:11677608
A:Accession: AB0707
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-310 <PAR>
A:Cross-references: UNIPARC:UP10000059DC1; GB:AL513382; PIDN:CADO2027.1; PID:JL6502864;
C:Genetic: STY1785
A:Gene: STY1785
Superfamily: 6-phosphofructokinase 2

A:Reference number: A91513; MWID:84262485; PMID:6535149
A:Accession: A24950
A:Molecule type: DNA
A:Residues: 1-25, 'ENCNAVPHRCSN',37-40,42-154, 'LRXNK',160, 'SAMSSTVL',169, 'QG',172-244, 'AL',
A:Cross-references: UNIPARC:PI01000016F3C7; GB:K02500; NID:g147151; PIDN:AAA24321.1; PID:6
A:Experimental source: strain K12
R:Daidal, F.
J. Mol. Biol. 168, 285-305, 1983
A:Title: Molecular cloning of the gene for phosphofructokinase-2 of *Escherichia coli* and
A:Reference number: 156378; MWID:83294514; PMID:6510120
A:Accession: 156378
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-25, 'ENCNAVPHRCSN',37-38 <RES>
A:Cross-references: UNIPARC:PI01000016F3C6; GB:K00128; NID:g147149; PIDN:AAA24320.1; PID:6
C:Comment: This enzyme is not to be confused with 6-phosphofructo-2-kinase (EC 2.7.1.105)
C:Genetics:
A:Gene: pfkF
A:Map position: 38 min
C:Complex: homotetramer
C:Function:
A:Description: catalyzes phosphorylation of fructose 6-phosphate to fructose 1,6-bisphosph
A:Pathway: glycolysis
C:Superfamily: 6-phosphofructokinase 2
A:Keywords: glycolysis, phosphotransferase

A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallet
 lech, J.; Harwood, C.R.; Henauf, A.; Hilbert, H.; Holtsappel, S.; Hosono, S.; Hullo, M.F.
 Koetter, P.; Konigstein, G.; Krogli, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
 A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel
 Y. M.; Ogawa, K.; Ogiwara, A.; Oudgas, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
 Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadie, Y.; Sato, T.; Scanlon,
 A;Authors: Schleich, S.; Schroefer, R.; Seiffert, P.; Sekiguchi, J.; Sekowska, A.; Serot
 akenchi, M.; Tamakoshi, A.; Tanaka, T.; Terpetra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
 T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
 A;Authors: Yoshikawa, H.F.; Zumein, E.; Yoshikawa, H.; Danchin, A.
 A;Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.
 A;Reference number: A69580; MUID:96044033; PMID:9384377
 A;Accession: A69627
 A;Status: Preliminary; nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-303 <KUN>
 A;Cross-references: UNIPROT:O31714; UNIPARC:UPI000006032D; GB:Z99111; GB:AL009126; NID:G
 A;Experimental source: strain 168
 C;Genetics:
 A;Gene: frub
 C;Superfamily: 6-phosphofructokinase 2

Query Match 17.5%; Score 300.5; DB 2; Length 303;
 Best Local Similarity 29.4%; Pred. No. 1.8e-14;
 Matches 98; Conservative 59; Mismatches 115; Indels 61; Gaps 14;
 QY 14 IITLNNPALDITTSVDVVRPTEKMRGCA---PRYD---PGGGGINVARIY---HVLGGC 64
 DB 2 IYVTLNP-----SVDIYVVEDEFTVGSLRRSYDTTPGGKGINVRLLKRRHV---A 52
 QY 65 STALFPAGSGTSLMLALLDAGVFPFVPIAASSTRESFTVN-ESRTAKQYRFLVPGESL 123
 DB 53 SKALFGVGGFTGEXIKTFLEBENLETARSEVKGDR---INVKLKTEDETEINGQGPFI 108
 QY 124 T-----VABEOCIDLKGAASAFVYASGSLPFGVAAADYQYQVADICRRSPFLILD 177
 DB 109 SDEDFKALFEQFQSQE-----GDIYVLAGSIPSSLPHTYEXIAACQQAARVVL 161
 QY 178 TSGGG-LOHISSGVFLKASVRELRECVSELTPEPQAAHELDIDGRALVVVSLGS 236
 DB 162 ISGGLALKATEMKRPFLLKPNHHELCMEGTATITSEVAVPFGKLVLEG-ABHYIVSMAG 220
 QY 237 QGALLATRHASHRPSIPW-TAVSGVAGADMAAIVTGLSRGWSLINSVRLGNAAAGAM 295
 DB 221 DGAALLFTBEAVY-FANVVKGLVNSVAGDSVAGFLAGISQKPLEBAFRLGVTSGSA- 278
 QY 296 LITPETAACNRDVERFPFLAABPTEVCQDQYV 328
 DB 279 -----TAFSEBELGTBEFV 291

RESULT 11
 D89841
 fructose-1-phosphate kinase [imported] - *Staphylococcus aureus* (strain N315)
 C;Species: *Staphylococcus aureus*
 C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
 C;Accession: D89841
 R;Kiroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc
 ma, A.; Mizutani-Uji, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
 C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiratazu, K.
 Lancet 357, 1225-1240, 2001
 A;Title: Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*.
 A;Reference number: A89758; MUID:21311952; PMID:11418146
 A;Accession: D89841
 A;Status: Preliminary
 A;Molecule type: DNA
 A;Residues: 1-306 <KUN>
 A;Cross-references: UNIPROT:O9KWK1; UNIPARC:UPI00000CA92D; GB:BA000018; PID:g13700590; F
 A;Experimental source: strain N315
 C;Genetics:
 A;Gene: frub
 C;Superfamily: 6-phosphofructokinase 2

Query Match 16.5%; Score 283; DB 2; Length 306;
 Best Local Similarity 26.6%; Pred. No. 3.4e-13;
 Matches 81; Conservative 66; Mismatches 139; Indels 18; Gaps 7;
 QY 14 IITLNNPALDITTSVDVVRPTEKMRGCAAPRYDGGGGINVARIYVAGCGSTALFPAGG 73
 DB 2 IYVTLNPSSIDYIYTFNDPFIKDLGNRATAYKRRAGGGINVRVLTLDVSTALTRAGG 61
 QY 74 STGSLMALLDGAGVFPFVPIAASSTRESFTVN-ESRTAKQYRFLVPGESLTVABEOCFL 132
 DB 62 FPGKFIDTLNNSAIQSNFIEVEDER---INVKLKTEGETEINAGPHITSTQFQQL 117
 QY 133 DELRGAASAFVYASGSLPFGVAAADYQYQVADICRRSPFLILDTSGGGLQHTSSGV-- 190
 DB 118 QQIKRN-TTSESDIYVAGSVSSISPSDAVQAQITPAQTAKLVDAE---KEIAESVLP 172
 QY 191 ---FLKASVRELRECVSELTPEPQAAHELDIDGRALVVVSLSGQALATRHAS 247
 DB 173 FHPFLFKNDLEFWMNTTVNSDPTVIKGRLLVDKG-AQSVYVSLGGGCAIYIDKEIS 231
 QY 248 HRFSSIPMTAVSGVAGADMAAIVTGLSRGWSLINSVRLGNAAAGAMLLTPGTACNRD 307
 DB 232 IKAVNPGKVVNTVYSGSDSTVAGVAGIAGLITRKAFQQAACGTATFDEDLA--TRD 289
 QY 308 DVER 311
 DB 290 AIEK 293

RESULT 12
 A11735
 fructose-1-phosphate kinase [imported] - *Listeria innocua* (strain Clp11262)
 C;Species: *Listeria innocua*
 C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
 C;Accession: A11735
 R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker
 .; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurgey, O.; Entian, K.D.; Feihn, H.
 D.; Jones, L.M.; Karet, U.
 Science 294, 849-852, 2001
 A;Authors: Krefit, J.; Kuhn, M.; Kunst, F.; Kurapkut, G.; Madueno, E.; Maitournam, A.; Mat
 ok, C.; Schlueder, T.; Simons, N.; Tietz, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
 A;Title: Comparative genomics of *Listeria* species.
 A;Reference number: A1077; MUID:21537279; PMID:11679669
 A;Accession: A11735
 A;Status: Preliminary
 A;Molecule type: DNA
 A;Residues: 1-307 <GUA>
 A;Cross-references: UNIPROT:Q928V2; UNIPARC:UPI00000CC878; GB:AL592022; PIDN:CA97657.1;
 A;Experimental source: strain Clp11262
 C;Genetics:
 A;Gene: frub
 C;Superfamily: 6-phosphofructokinase 2

Query Match 16.0%; Score 275; DB 2; Length 307;
 Best Local Similarity 27.3%; Pred. No. 1.3e-12;
 Matches 77; Conservative 58; Mismatches 141; Indels 6; Gaps 4;
 QY 14 IITLNNPALDITTSVDVVRPTEKMRGCAAPRYDGGGGINVARIYVAGCGSTALFPAGG 73
 DB 2 IYVTLNPSSIDYIYQIOLKLGELNRRKQDYKLPGGGINVRVLTLDVSTALTRAGG 61
 QY 74 STGSLMALLDGAGVFPFVPIAASSTRESFTVNESRTAKQYRFLVPGESLTVABEOCFLD 133
 DB 62 FTGGFIDWLKNEBIOGFYIVDDTRINIKLKHGETEINGL--GPAISKEIKRFL- 117
 QY 134 ELRGAASAFVYASGSLPFGVAAADYQYQVADICRRSPFLILDTSGGG-LQHSISGVFL 192
 DB 118 EVMDDKTVAGDIYVILISGVPSPSLGNDPFYDKIIOICKEKQAQFMIDTTCQELIDALPNRPIL 177
 QY 193 LKASVRELRECVSELTPEPQAAHELDIDGRALVVVSLSGQALATRHASHRFS 252
 DB 178 IKPNHHELELFVGLKDSVEDLIPYKKCLELG-AQHIVVSMAGDALFTGDEVPADA 236

Job time : 14.4872 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleation Ltd.

OM protein - protein search, using sw model

Run on: March 23, 2006, 04:56:43 ; Search time 86.7035 Seconds
(without alignments)
2758.529 Million cell updates/sec

Title: US-10-617-038-22
Perfect score: 1720
Sequence: 1 MTEPAAWDEGKPRITITLTMN.....TEVGQDQYVHPPIVPEASP 339

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 segs, 705528306 residues
Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1720	100.0	339	Q7TZ43_MYCBO	Q7TZ43_mycobacteri
2	1720	100.0	339	O86352_MYCTU	O86352_mycobacteri
3	749	43.5	308	Q4J5V5_AZOV1	Q4J5V5_azobacter
4	667	38.8	347	Q4NK08_9M1CC	Q4NK08_9M1CC
5	648	37.7	312	Q6KCU0_RHOPA	Q6KCU0_rhodopseudo
6	603.5	35.1	313	Q631V9_BURPS	Q631V9_burkholderi
7	597.5	34.7	313	Q62EC2_BURMA	Q62EC2_burkholderi
8	574	33.4	295	O89189_BRAJA	O89189_bradyniob
9	539.5	31.4	310	O57PV8_SALCH	O57PV8_salmonella
10	539.5	31.4	310	O8ZPT5_SALTY	O8ZPT5_salmonella
11	534.5	31.1	310	O8Z6H7_SALTI	O8Z6H7_salmonella
12	534.5	31.1	310	O5PH98_SALPA	O5PH98_salmonella
13	531.5	30.9	309	K6FPF2_ECOLI	P06999_escherichia
14	530.5	30.8	309	O83140_SHIFL	O83140_shigella fl
15	530.5	30.8	309	O8XE22_ECO57	O8XE22_escherichia
16	524.5	30.5	310	O8FH20_ECOL6	O8FH20_escherichia
17	355	20.6	317	O8ZH57_STRAW	O8ZH57_streptomyce
18	355	20.6	326	Q4NDB8_9M1CC	Q4NDB8_arthrobacte
19	354	20.6	307	Q73A75_BACCI	Q73A75_bacillus ce
20	354	20.6	310	O9KXV8_STRCO	O9KXV8_streptomyce
21	349	20.3	315	O9KXV8_STRCO	O9KXV8_streptomyce
22	348	20.2	307	O63CY0_BACIZ	O63CY0_bacillus ce
23	343.5	20.0	323	Q67Q12_STWTH	Q67Q12_arthrobacte
24	334	19.4	332	Q4NKX2_9M1CC	Q4NKX2_arthrobacte
25	330	19.2	309	O8ZGF5_STRAW	O8ZGF5_streptomyce
26	328	19.1	303	O8ES06_OCEIH	O8ES06_oceanobacil
27	327.5	19.0	326	Q4H6K1_9B8FO	Q4H6K1_enterococcu
28	322	18.7	313	O833W9_ENTFA	O833W9_enterococcu
29	320	18.6	321	O67RB4_STWTH	O67RB4_syndiodacte
30	318.5	18.5	315	O9ZRP6_DEIRA	O9ZRP6_deinococcus
31	316	18.4	292	O6AD91_LEIRXX	O6AD91_leifsonia x

32	314	18.3	303	2	Q73B5_BACCI	Q73B5_bacillus ce
33	312.5	18.2	315	2	O6AG92_LEIRX	O6AG92_leifsonia x
34	312	18.1	303	2	O6HP93_BACHK	O6HP93_bacillus th
35	312	18.1	303	2	O636V8_BACIZ	O636V8_bacillus ce
36	311	18.1	303	2	O81WV7_BACAN	O81WV7_bacillus an
37	309.5	18.0	308	2	O97F04_CLOAB	O97F04_clostridium
38	309.5	18.0	316	2	O5YV79_NOCRA	O5YV79_nocardia fa
39	308	17.9	303	2	O5K1W2_GEOKA	O5K1W2_geobacillus
40	307	17.8	310	1	LACC_STRMU	P26421_streptococc
41	305	17.7	303	2	Q4MJY4_BACCE	Q4MJY4_bacillus ce
42	304	17.7	309	2	O97MG6_CLOAB	O97MG6_clostridium
43	304	17.7	315	2	O88PQ4_PSEPK	O88PQ4_pseudomonas
44	303	17.6	324	2	O8PFP0_CORER	O8PFP0_corynebacte
45	300.5	17.5	300	2	O6Y9B0_BORHE	O6Y9B0_borrelia ne

ALIGNMENTS

RESULT 1	Q7TZ43_MYCBO	PRELIMINARY;	PRT;	339 AA.
ID	Q7TZ43			
AC	Q7TZ43			
DT	01-OCT-2003 (TREMBLrel. 25, Created)			
DT	01-OCT-2003 (TREMBLrel. 25, Last sequence update)			
DT	01-MAR-2004 (TREMBLrel. 26, Last annotation update)			
DE	Probable phosphofructokinase pfkb (PHOSPHOKINASE)			
DE	(EC 2.7.1.-)			
GN	Name=pfkb; OrderedlocusNames=Mb2054c;			
OS	Mycobacterium bovis;			
OC	Bacteria; Actinobacteria; Actinomycetales;			
OC	Corynebacteriaceae; Mycobacteriaceae; Mycobacterium;			
OC	Mycobacterium tuberculosis complex.			
OX	NCBI_TaxID=1765;			
RN	[1]			
RP	NUCLEOTIDE SEQUENCE.			
RC	STRAIN=AF2122/97;			
RX	MEDLINE=22709107; PubMed=12788972; DOI=10.1073/pnas.1130426100;			
RA	Garnier T., Sigmeleer K., Camus J.-C., Medina N., Mansoor H.,			
RA	Pryor M., Duhoy S., Gordin S., Lacroix C., Monsemp C., Simon S.,			
RA	Harris B., Actin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,			
RA	Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;			
RT	"The complete genome sequence of Mycobacterium bovis."			
RL	Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).			
DR	EMBL; BX248341; CAD96907.1; -; Genomic DNA.			
DR	GO; GO:0016301; F:kinase activity; IEA.			
DR	GO; GO:0016740; F:transferase activity; IEA.			
DR	InterPro; IPR002173; PKB.			
DR	InterPro; IPR01611; PKB_region.			
DR	Pfam; PF00294; PKB_1.			
DR	PROSITE; PS00583; PKB_KINASES_1; 1.			
KW	Complete proteome; Kinase; Transferase.			
SO	SEQUENCE 339 AA; D7735363FF73DAD0 CRC64;			
Query Match	100.0%; Score 1720; DB 2; Length 339;			
Best Local Similarity	100.0%; Pred. No. 2, 8e-109;			
Matches	339; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1	MTEPAAWDEGKPRITITLTMN	PALDITTSVDVVRTEKRCGAPRYDGGGINARIVHV	60
DB	1	MTEPAAWDEGKPRITITLTMN	PALDITTSVDVVRTEKRCGAPRYDGGGINARIVHV	60
QY	61	LGGSSTLFPAGSGTSLMAL	LGADAGVPRVPIAATSTRESFTVNSRTAKOYRFVLP	120
DB	61	LGGSSTLFPAGSGTSLMAL	LGADAGVPRVPIAATSTRESFTVNSRTAKOYRFVLP	120
QY	121	PSLTVAQEQCLDELRCGAASAFVVASGSLPPGVADYYQ	QRVADICRSSTPLILDTSG	180
DB	121	PSLTVAQEQCLDELRCGAASAFVVASGSLPPGVADYYQ	QRVADICRSSTPLILDTSG	180
QY	181	GGLGHISSGVFLAKASVRELRCEVSELTPEQ	LAAHLLIDGRANVVVVSIGSGAL	240
DB	181	GGLGHISSGVFLAKASVRELRCEVSELTPEQ	LAAHLLIDGRANVVVVSIGSGAL	240

QY 241 LATERASHRFSIPMTAVSGVAGADAMVAATVGLSRGWSLIKSVRLGNAAGAMLLTPG 300
 DB 241 LATERASHRFSIPMTAVSGVAGADAMVAATVGLSRGWSLIKSVRLGNAAGAMLLTPG 300
 QY 301 TAAACNRDDVERFFELAAEPTEVGODQYVWHPIVNPASP 339
 DB 301 TAAACNRDDVERFFELAAEPTEVGODQYVWHPIVNPASP 339

RESULT 2
 086352_MYCTU PRELIMINARY; PRT; 339 AA.
 ID 086352_MYCTU PRELIMINARY; PRT; 339 AA.
 AC 086352; Q7D7L4;
 DT 01-NOV-1998 (TREMblrel. 08, Created)
 DT 01-NOV-1998 (TREMblrel. 08, Last sequence update)
 DT 13-SEP-2005 (TREMblrel. 31, Last annotation update)
 DE Probable phosphotransferase PfkB (PHOSPHOTRANSFERASE) (EC 2.7.1.-)
 GN Name=PfkB; OrderedLocustNames=MT2088, Rv2023c;
 OS Mycobacterium tuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteriales; Actinomycetales;
 OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium;
 OC Mycobacterium tuberculosis complex.
 CX NCBI_TaxID=1773;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=H37Rv;
 RX MEDLINE=98295987; PubMed=9634230; DOI=10.1038/31159;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C.M.,
 RA Harris D.E., Gordon S.V., Eigmeier K., Gas S., Barry C.E. III,
 RA Tekala F., Badcock K., Basham D., Brown D., Chillingworth T.,
 RA Connor R., Davies R.M., Devlin K., Feltwell T., Gentles S., Hamlin N.,
 RA Holroyd S., Hornsby T., Jagels K., Krogh A., McLean J., Moule S.,
 RA Murphy L.D., Oliver S., Osborne J., Quail M.A., Rajandream M.A.,
 RA Rogers J.E., Rutter S., Seeger K., Skelton S., Squares R.,
 RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
 RT "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence."
 RL Nature 393:537-544(1998).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=CDC 1551 / Oshkosh;
 RX MEDLINE=22206494; PubMed=12218036;
 RX DOI=10.1128/JB.184.19.5479-5490.2002;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J.D., DeBoy R.T., Dodson R.J., Gwinn M.L., Haft D.H.,
 RA Hickey E.K., Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D.,
 RA Salzberg S.L., Delcher A.L., Utermack T.R., Weidman J.F., Khouri H.M.,
 RA Gill J., Mikula A., Bishai W., Jacobs W.R. Jr., Venter J.C.,
 RA Fraser C.M.;
 RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains."
 RL J. Bacteriol. 184:5479-5490(2002).
 DR EMBL; BX842578; CA17243.1; -; Genomic DNA.
 DR EMBL; AE000516; AA64367.1; -; Genomic DNA.
 DR PIR; D70942; D70942.
 DR TIGR; MT2088; -;
 DR TubercuLab; Rv2023c; -;
 DR GO; GO:0016301; F:kinase activity; IEA.
 DR InterPro; IPR02173; PFKB.
 DR InterPro; IPR01611; PFKB_region.
 DR Pfam; PF00294; PFKB_1.
 DR PROSITE; PS00583; PFKB_KINASES_1; 1.
 KW Complete proteome; Kinase; Transferase.
 SQ SEQUENCE 339 AA; 35401 MW; D773E363F73DAD0 CRC64;

Query Match 100.0%; Score 1720; DB 2; Length 339;
 Best Local Similarity 100.0%; Pred. No. 2.8e-109;
 Matches 339; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTEPAMDEGKPRITLTMPALDITTSVDVVRPTEKRCGAPRYDGGGINVARIHV 60
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||

DB 1 MTEPAMDEGKPRITLTMPALDITTSVDVVRPTEKRCGAPRYDGGGINVARIHV 60
 QY 61 LGGCSTALFPAGSSTGSLMALLDGADVPRVPIPIAATSTESTVNESTAKQRYFVLP 120
 DB 61 LGGCSTALFPAGSSTGSLMALLDGADVPRVPIPIAATSTESTVNESTAKQRYFVLP 120
 QY 121 PSLTVAEQECLDELGAASAAFPVAGSGLPGVADYQRAVDICRSTPLILDTSG 180
 DB 121 PSLTVAEQECLDELGAASAAFPVAGSGLPGVADYQRAVDICRSTPLILDTSG 180
 QY 181 GGLQHTSSGVFLKASVRELRECVSELTTEPOLAAHEILRGRAEVVVSIGSOGAL 240
 DB 181 GGLQHTSSGVFLKASVRELRECVSELTTEPOLAAHEILRGRAEVVVSIGSOGAL 240
 QY 241 LATERASHRFSIPMTAVSGVAGADAMVAATVGLSRGWSLIKSVRLGNAAGAMLLTPG 300
 DB 241 LATERASHRFSIPMTAVSGVAGADAMVAATVGLSRGWSLIKSVRLGNAAGAMLLTPG 300
 QY 301 TAAACNRDDVERFFELAAEPTEVGODQYVWHPIVNPASP 339
 DB 301 TAAACNRDDVERFFELAAEPTEVGODQYVWHPIVNPASP 339

RESULT 3
 0405V5_AZOVI PRELIMINARY; PRT; 308 AA.
 ID 0405V5_AZOVI PRELIMINARY; PRT; 308 AA.
 AC 0405V5;
 DT 13-SEP-2005 (TREMblrel. 31, Created)
 DT 13-SEP-2005 (TREMblrel. 31, Last sequence update)
 DT 13-SEP-2005 (TREMblrel. 31, Last annotation update)
 DE Carboxylate kinase, PfkB.
 GN ORFNames=AvindRAFT_5718;
 OS Azotobacter vinelandii AZOP.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Azotobacter.
 CX NCBI_TaxID=322710;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=AZOP;
 RG US DOE Joint Genome Institute (JGI-PGF);
 RA Copeland A., Lucas S., Lapidus A., Barry K., Dettler C., Glavina T.,
 RA Hammon N., Israni S., Pitluck S., Richardson P.;
 RT "Sequencing of the draft genome assembly of Azotobacter vinelandii
 RT AZOP."
 RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=AZOP;
 RG US DOE Joint Genome Institute (JGI-ORNL);
 RA Larimer F., Land M.;
 RT "Annotation of the draft genome assembly of Azotobacter vinelandii
 RT AZOP."
 RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=AZOP;
 RA DOE Joint Genome Institute;
 RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=AZOP;
 RG US DOE Joint Genome Institute (JGI-PGF);
 RA Copeland A., Lucas S., Lapidus A., Barry K., Dettler C., Glavina T.,
 RA Hammon N., Israni S., Pitluck S., Richardson P.;
 RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
 CC -1- CAUTION: The sequence shown here is derived from an
 CC preliminary data.
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 DR EMBL; AAU03000001; EAM07036.1; -; Genomic DNA.
 KW Kinase.
 SQ SEQUENCE 308 AA; 32503 MW; 035PD99DBC74190D CRC64;

Query Match 43.5%; Score 749; DB 2; Length 308;


```

Qy 193 KASVRELBECVGESELTTEPEOLAAHHELIDRGRAEWWWVSLGSGOALLATPRASHRFS 252
Db 184 IKPNIREMOELAGRPLGCEKEWDDARAHITYGQVEIVALSILGHLVLTIKOEARSPA 243
Qy 253 IPMTAVSGGAGADAMVAATTVGSGRMSLISKVRLNMAAGAAMLTPTGNAACNRDVERF 312
Db 244 LPVTIVSTVGAODSFLAAITAHRLAOGGSVADAFRLGLAAGSATLTVHGTOLCGPAVERL 303
Qy 313 F-ELAAP 319
Db 304 YGQVITIEP 311

```

RESULT 6	
Q631V9_BURPS	
ID Q631V9_BURPS	PRELIMINARY;
	PRT; 313 AA.

DT 25-OCT-2004 (TREMBlrel. 28, Created)
 DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)
 DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)
 DE Putative 6-phosphotransferase.
 GN Ordered locus names=BPSS1957;
 OS Burkholderia pseudomallei (pseudomonas pseudomallei).
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Burkholderiaceae; Burkholderia; pseudomallei group.
 OX NCBI_TaxID=28450;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=K96243;
 RX PubMed15377794; DOI=10.1073/pnas.040302101;
 RA Holden M.T.G., Tilball R.W., Peacock S.J., Cerdeno-Tarraga A.-M.,
 RA Atkins T., Crossman L.C., Pitt T., Churcher C., Mungall K.L.,
 RA Bentley S.D., Seabaith M., Thomson N.R., Baoun N., Beacham I.R.,
 RA Brooks K., Brown K.A., Brown N.F., Challis G.L., Cherevach I.,
 RA Chillingworth T., Cronin A., Crosscut B., Davis P., Deshazer D.,
 RA Fellwell T., Frazer A., Hance Z., Hauser H., Holtroyd S., Jagels K.,
 RA Keith K.E., Maddison M., Moule S., Price C., Quail M.A.,
 RA Rabinowitch E., Rutherford K., Sanders M., Simmonds M.,
 RA Songer-Smith S., Stevens K., Tunnacliff S., Veenstra-Thavet M.,
 RA Whitehead S., Yeats C., Barrett B.G., Overton P.C.F., Parkhill J.;
 RT "Genomic plasticity of the causative agent of melioidosis,
 RT Burkholderia pseudomallei".
 RL Proc. Natl. Acad. Sci. U.S.A. 101:14240-14245 (2004).
 DR EMBL; BX571966; CAH39435.1; -, Genomic_DNA.
 DR GO; GO:0016301; P:kinase activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR InterPro; IPR002173; PfkB.
 DR InterPro; IPR01611; PfkB_region.
 DR Pfam; PF00294; PfkB; 1.
 DR PROSITE; PS00583; PFKB_KINASES_1; 1.
 KW Complete proteome; Kinase.
 QO SEQUENCE 313 AA; 31941 MW; 2CE607B495039439 CRC64;

Query Match	35.1%	Score 603.5;	DB 2;	Length 313;
Best Local Similarity	42.2%;	Pred. No. 3.9e-33;		
Matches 128:	Conservative 59;	Mismatches 115;	Indels 1	

QY 12 PRILITLWNPALDITSDVVAAPTEMRGCAPIYDGGGGINARIVHVLGGSTLAFPA 71

Db 2 PEIVYLTPNPALDVATCVERVYDTRKLRGCGPARRDPGGGGINARIVLTRLGADCSAVYLA 61

QY 72 GGSSTGLMALIGDAGVPRFVPIAASTRESFTVMSERTAKOYRFVLPGBSLTVAAEOQC 131

Db 62 GGGTGLALRGLIADBSGVRAHGIIDINGETRENTSVLETSRGERRFVLPGBALAAHMPRC 121

QY 132 LDELKGAASAAPFVVAAGSLPRPGVAADYYQRYVADICRBSSTPLILDTSGGCI-QHITSSEY 190

Db 122 VEAIGRLADASRYLYWMSGSLPRGMPEDDCYARLARASARGRVTVVDTSGPALAAALDAGY 181

QY 191 FLUKASVVELRECVSESLTEPEQOLAAAEHLDRGAARVWVWSLGSQGHALLTRHSHRP 250

Db 182 YLVKRSIQLDELRLVTLPLEDDGARLLAAAVIAGGAQIVALLTLGAGALVWSRDDAVRL 241

QY 251 SSIPPTAAVSGVAGGAMVAALITVGSLSRGSLLIKSVLGNAGAAMLITPGTAACNRDVE 310
 ||||| : | : | : | : | : | : | : | :
Dd 242 PGVKAVVASAIGAGDSFVAGLVVALNRRGANVADAAARHALLAAAASLSLTGTALGTREDLA 301
 ||| : | : | : | : | : | : | : | :
QY 311 RFF 313
 || :
Dd 302 RIY 304

RESULT 7	
Q62EC2 BURMA	
ID Q62EC2_BURMA PRELIMINARY;	PRT; 313 AA

DR 25-OCT-2004 (TREMBLrel. 28, Created)
DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)
DE 6-phosphofructokinase (EC 2.7.1.11).
GN Name=pfk; OrderedLocNames=BMAA0117;
OS Burkholderia mallei (Pseudomonas mallei).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia.
OX NCBI_TaxID=13373;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 23344;
RX PubMed=15377793; DOI=10.1073/pnas.0403306101;
RA Nieman W.C., Deshazer D., Kim H.S., Tetteelin H., Nelson K.E.,
RA Feldblyum T.V., Ulrich R.L., Roming C.M., Brintac L.M.,
RA Daugherty S.C., Davidsen T.D., Deboy R.T., Dimitrov G., Dodson R.J.,
RA Durkin A.S., Gwim M.L., Hafé D.H., Khouri H.M., Kolonay J.F.,
RA Madupu R., Mohammoud Y., Nelson W.C., Radeue D., Romero C.W.,
RA Saefer N., Zhou L., Fraser C.M.,
RA "Structural flexibility in the Burkholderia mallei genome.";
RT Proc. Natl. Acad. Sci. U.S.A. 101:14246-14251(2004).
RL EMBL: CP000011; AAU46348.1; -; Genomic_DNA.
DR TIGR: BMAA0117; -
DR GO: GO:0003872; F:6-phosphofructokinase activity; IEA.
DR GO: GO:0016301; F:kinase activity; IEA.
DR GO: GO:0016740; F:transferase activity; IEA.
DR InterPro: IPR002173; PFKB.
DR InterPro: IPR011611; PFKB_region.
DR Pfam: PF00294; PFKB; 1.
DR PROSITE: PS00583; PFKB_KINASES_1; 1.
DR Complete proteome; Kinase; Transferase;
SQ SEQUENCE 313 AA; 2CB607B495045FP9 CRC64;

Query Match	34.7%	Score 597.5;	DB 2;	Length 313;
Best Local Similarity	41.9%	Pred. No. 1e-32;		
Matches 127; Conservative	59;	Mismatches 116;	Indels 1;	Gaps 1

```
QY      12 P R I I T N N P A D I T T S V D V A R P T E K M C G A R Y P P G G G I N V A I V H L G C S T L F P A    71
       : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      2 P E I Y I T N P A I D V A T C V E R Y I D T R K L C G P A R R P G G G I N V A V I T L G A D C A S V I L A    61

QY      72 G G S T S L L M A L L G D A V F P R V I P I A S T R E S E T V N E S R T A K O Y R F V L P G S I T V A E Q E O C   131
Db      62 G G G T G L A R G L L A D E G V B A H G I D I G E T R E N F S V L E T S T G R E F R F V L P G P A L A A H E M P R C   121

QY      132 L D E L R G A A S A A F V V A S G S L P P G V A A D Y O R A D I C R S S T F L I D T S G G L - Q H I S S E V    190
Db      122 V E A L G R L D A S R Y I L W M S G S L P P G M D D C Y A R L A R P A S A R G V R T V D T S G P A L A A A L D A C V    181

QY      191 F L U K A V R E L R E C V S E L L T B E O C L A A A H E L I D R G A E V V V V S I Q G A L L T R H A S H R F     250
Db      182 Y L V P S L G E L P A L T L P L E D D C A R L A A R A Y A G S A Q V A L T L G D A G L V V S R D D A V L     241

QY      251 S S I P T A V A S G V G A G D A M V A A I T V G L S R G W S L I K S Y R L G N A A G A M L T P G T A A C N R D V E    310
Db      242 P G V K A V A R S A I G A G D S F V A G V A A L N R G A N V A D A A R H L A A S A S I L S T G P A L G T M E D I A    301

QY      311 R F F 313
       :
```

Db 302 R1Y 304

RESULT 8
089189_BRAJA PRELIMINARY; PRT; 295 AA.

AC 089189;
ID 089189_BRAJA PRELIMINARY; PRT; 295 AA.

DT 01-JUN-2003 (Tremblrel. 24, Created)
DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE PF6B family carbohydrate kinase.
GN OrderedLocustNames=blx14659;
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
ON NCBI_TaxID=375;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=USDA 110;
RX MEDLINE=2248498; PubMed=12597275;
RA Kaneo T., Nakamura Y., Sato S., Minamisawa K., Uchiyama T.,
RA Sasamoto S., Watanabe A., Idegawa K., Iriuchida M., Kawashima K.,
RA Kohara M., Matsumoto M., Shimpo S., Tsunoka H., Wada T., Yamada M.,
RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
Bradyrhizobium japonicum USDA110."
RL DNA Res. 9:189-197(2002).
DR EMBL: BA000040; BAC49924.1; -: Genomic DNA.
DR GO: GO:0016301; F:kinase activity; IEA.
DR InterPro: IPR002173; PF6B.
DR InterPro: IPR011611; PF6B_region.
DR Pfam: PF00294; PF6B_1.
DR PROSITE: PS00583; PF6B_KINASES_1; 1.
DR PROSITE: PS00584; PF6B_KINASES_2; UNOWN_1.
KW Complete proteome; Kinase.
SQ SEQUENCE 295 AA; 30724 MW; E71C73B9308B3455 CRC64;

Query Match 33.4%; Score 574; DB 2; Length 295;
Best Local Similarity 45.0%; Pred. No. 3.8e-31;
Matches 127; Conservative 47; Mismatches 100; Indels 8; Gaps 3;

QY 34 PTEKRCGAPRYDGGGGINVARIYHVGCGSTALFPFGSGTSLMALLDAGVPPRYI 93
DB 3 PYTEKRCGAPRYDGGGGINVARIYHVGCGSTALFPFGSGTSLMALLDAGVPPRYI 62

QY 94 PIAASTRESFTVNSRTAKOYRFLVGPSTLVAEQECLDELRGAASAFAVVASGSLRP 153
DB 63 PISNDTRDITITFDTSKEQFLVPGACLNPEWQCLDAIARLSPQAAFLVASGSLRP 122

QY 154 GVAADYQORVADICRSST---PLIDTSGGGLQ-HISSGVFLKASVRELRCVSGEL 209
DB 123 GAPADYQORVADICRSST---FASTGAAKVVDISGASIKAAELAGVYLKPNRREFQELAGISGA 178

QY 210 TEPBQLAAHELDIRGRAEVVVSIGSGALLATRHASHRPSSIPMTAVSGVGADMANVA 269
DB 179 DESSLEKGRLLIRYRIEILALISWPGGALLTRFDILRANGPLLEPVSVSAGDSFLG 238

QY 270 AITVGLSGWSLISVRLGNAAAMLLTPGTAACNRDVR 311
DB 239 AMVSRLANGDKLDSRLRYGVAGSAAALSPGTGLGADVHR 280

RESULT 9
057PV8_SALCH PRELIMINARY; PRT; 310 AA.

AC 057PV8;
DT 10-MAY-2005 (Tremblrel. 30, Created)
DT 10-MAY-2005 (Tremblrel. 30, Last sequence update)
DT 10-MAY-2005 (Tremblrel. 30, Last annotation update)
DE 6-phosphofructokinase II.
GN Name=pfkfb; OrderedLocustNames=SC1347;
OS Salmonella cholerae-suis (Salmonella enterica).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Salmonella.
ON NCBI_TaxID=591;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=SC-867;
RX PubMed=15781495;
RA Chiu C.-H., Tang P., Chu C., Hu S., Bao Q., Yu J., Chou Y.-Y.,
RA Wang H.-S., Lee Y.-S.;
RT "The genome sequence of Salmonella enterica serovar Choleraesuis, a
highly invasive and resistant zoonotic pathogen."
RL Nucleic Acids Res. 33:1690-1698(2005).
DR EMBL: AE017220; AAK65253.1; -: Genomic DNA.
KW Complete proteome; Kinase.
SQ SEQUENCE 310 AA; 32565 MW; 38C7B972571F3F CRC64;

Query Match 31.4%; Score 539.5; DB 2; Length 310;
Best Local Similarity 41.8%; Pred. No. 9e-29;
Matches 127; Conservative 54; Mismatches 118; Indels 5; Gaps 4;

QY 13 RIITLTNMPALDITTSVDVVRPTEKRCGAPRYDGGGGINVARIYHVGCGSTALFPAG 72
DB 3 RIYTLTAPSDSATITPQIYBEGKLRCSAPVFEPGGGGINVARIAHLLGGTATATPAG 62

QY 73 GSTGSLMALIGDGVPPRYIPIASTRESFTVNSRTAKOYRFLVGPSTLVAEQECL 132
DB 63 GATEHVLVALLDENVSTVDADKDWTRONLHVHVESSGQRYRPMGCATLDDDEFRO-L 121

QY 133 DELGAASAFAVVASGSLPPGVADYQORVADICRSSTPLIDTSGGGLQHSIS--GV 190
DB 122 EEOYLEIESGAILVSSLPGVVEKLTQLISAQOGKRCITDSSGDALITALAGDI 181

QY 191 FLKASVRELRCVSGELTEPEQL-AAHELDIRGRAEVVVSIGSGALLATRHASHR 249
DB 182 ELVKNPKELSAVLNVRD-LTPQDDVRKAQELVSGARRVVSILGPGALIDISENCIO 240

QY 250 FSSIPMTAVSGVGADMANVAITVGLSRGWSLISVRLGNAAAMLLTPGTAACNRDDV 309
DB 241 VVPEPVVSOSTVGAGDSMGAMTLKLAQDASLEBMYRFGVAAGSAATLNOGTBLCSHDT 300

QY 310 ERFF 313
DB 301 QKIT 304

RESULT 10
08ZPT5_SALTY PRELIMINARY; PRT; 310 AA.

AC 08ZPT5;
DT 01-MAR-2002 (Tremblrel. 20, Created)
DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
DE 6-phosphofructokinase II (EC 2.7.1.11).
GN Name=pfkb; OrderedLocustNames=STM1326;
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
ON NCBI_TaxID=602;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=L72;
RX MEDLINE=21534948; PubMed=11677609; DOI=10.1038/35101614;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Portolillo S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Flores L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
L72."
RL Nature 413:852-856(2001).
DR EMBL: AE008757; AL20251.1; -: Genomic DNA.
DR GO: GO:0003872; F:6-phosphofructokinase activity; IEA.
DR GO: GO:0016301; F:kinase activity; IEA.
DR GO: GO:0016740; F:transferase activity; IEA.

DR GO:0005975; P:carbohydrate metabolism; IEA.
 DR InterPro; IPR002173; Pfam.
 DR InterPro; IPR011611; Pfam_region.
 DR Pfam; PF00294; Pfam; 1.
 DR PROSITE; PS00583; PFKB_KINASES_1; 1.
 DR PROSITE; PS00584; PFKB_KINASES_2; 1.
 DR Complete proteome; Kinase; transferase.
 KW SEQUENCE 310 AA; 32565 MW; 38CCTEB972571F3F CRC64;

Query March 31.4%; Score 539.5; DB 2; Length 310;
 Best Local Similarity 41.8%; Pred. No. 9e-29;
 Matches 127; Conservative 54; Mismatches 118; Indels 5; Gaps 4;

QY 13 RIITLTMNPALDITTSVDVVRPTEKRCGAPRYDPGGGGINVARIHVHVGCGSTALFPAG 72
 3 RIVTLTAPSLDASATITPQIYPEGKLRCSAPFEPGGGGINVARAIHAGTATATFPAG 62
 QY 73 GSTGSLMALIGDAGVPRVPIAASRSTVESRTAKQRYVLPGPSLTVAEQDCL 132
 63 GATGEHLVALADENVPSTVDAKDWTQNLHVHVSSEGYRFPVFGATLDDDEFQ-L 121
 QY 133 DELRGAASAAPVVAAGSLPPGVAADYQYRVADICRRSSTPLIDTSGGGLQHTSS--GV 190
 122 EBOVLEIESGAILIVISGSLPPGVKVEKLTQLISAQKQGRICIIDSSGDLTALALGDI 181
 QY 191 FLTKASVRELRECVSELTPEQL-AAAHELIDRGAEVVVVSLSGGALLATRHASHR 249
 182 ELVKPNKELSAVLVND-LTPDDVRKAQELVSGKARVVVSLGPGALGIDSENCIQ 240
 QY 250 FSSIPMTAVSGVAGDAMVAITVGLSRGMSLISVRLGNAAAGAMLLTPTAACRNDV 309
 241 VVPPVKSQSTVGAGDSMGVAMTLKLAQDASLEEMVRFVGAASATLNGTRLCSHDDT 300
 Db 310 ERF 313
 QY 301 QKTY 304

RESULT 11

0826H7_SALTI PRELIMINARY; PRT; 310 AA.

AC 0826H7_07CA82;
 DT 01-MAR-2002 (TREMBlrel. 20, Created)
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
 DT 10-MAY-2005 (TREMBlrel. 30, Last annotation update)
 DE 6-phosphofructokinase isozyme.
 GN Name=pfk; OrderedlocusNames=STY1785, t1206;
 OS Salmonella typhi.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Salmonella.
 OX NCBI_Taxid=601;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=CT18;
 RX MEDLINE=21534947; PubMed=11677608; DOI=10.1038/35101607;
 RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
 Churcher C.M., Mungall K.O., Bentley S.D., Holden M.T.G., Sebahia M.,
 Baker S., Basham D., Brooks K., Chillingworth T., Conerton P.,
 Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
 Feltham T., Hamlin N., Haeghe A., Hien T.T., Holtroyd S., Jørgen K.,
 Krogan A., Larsen T.S., Leach S., Mout R., O'Gaora P., Parry C.,
 Quail M.A., Rutherford K.M., Simmonds M., Skelton J., Stevens K.,
 Whitehead S., Barrall B.G.;
 RT "Complete genome sequence of a multiple drug resistant Salmonella
 enterica serovar Typhi CT18.";
 RL Nature 413:848-852(2001).
 RU [2]

RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=TY2 / ATCC 700931;
 RX MEDLINE=2531367; PubMed=12644504;
 RX DOI=10.1128/JB.185.7.2330-2337.2003;
 RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
 Burdand V., Kodyanski V., Schwartz D.C., Blattner F.R.;

RT "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
 and CT18.";
 RL J. Bacteriol. 185:2330-2337(2003).
 DR EMBL; AL627271; CA02027.1; -; Genomic DNA.
 DR EMBL; AB016838; AAO68862.1; -; Genomic DNA.
 DR GO:0016301; F:kinase activity; IEA.
 DR GO:0016740; F:transferase activity; IEA.
 DR GO:0005975; P:carbohydrate metabolism; IEA.
 DR InterPro; IPR002173; PFKB.
 DR InterPro; IPR011611; PFKB_region.
 DR Pfam; PF00294; PFKB; 1.
 DR PROSITE; PS00583; PFKB_KINASES_1; 1.
 DR PROSITE; PS00584; PFKB_KINASES_2; 1.
 KW Complete proteome; Kinase.
 KW SEQUENCE 310 AA; 32546 MW; 38CCTEB9724BBD3F CRC64;

Query March 31.4%; Score 534.5; DB 2; Length 310;
 Best Local Similarity 41.4%; Pred. No. 2e-28;
 Matches 126; Conservative 54; Mismatches 119; Indels 5; Gaps 4;

QY 13 RIITLTMNPALDITTSVDVVRPTEKRCGAPRYDPGGGGINVARIHVHVGCGSTALFPAG 72
 3 RIVTLTAPSLDASATITPQIYPEGKLRCSAPFEPGGGGINVARAIHAGTATATFPAG 62
 QY 73 GSTGSLMALIGDAGVPRVPIAASRSTVESRTAKQRYVLPGPSLTVAEQDCL 132
 63 GATGEHLVALADENVPSTVDAKDWTQNLHVHVSSEGYRFPVFGATLDDDEFQ-L 121
 QY 133 DELRGAASAAPVVAAGSLPPGVAADYQYRVADICRRSSTPLIDTSGGGLQHTSS--GV 190
 122 EBOVLEIESGAILIVISGSLPPGVKVEKLTQLISAQKQGRICIIDSSGDLTALALGDI 181
 QY 191 FLTKASVRELRECVSELTPEQL-AAAHELIDRGAEVVVVSLSGGALLATRHASHR 249
 182 ELVKPNKELSAVLVND-LTPDDVRKAQELVSGKARVVVSLGPGALGIDSENCIQ 240
 QY 250 FSSIPMTAVSGVAGDAMVAITVGLSRGMSLISVRLGNAAAGAMLLTPTAACRNDV 309
 241 VVPPVKSQSTVGAGDSMGVAMTLKLAQDASLEEMVRFVGAASATLNGTRLCSHDDT 300
 Db 310 ERF 313
 QY 301 QKTY 304

RESULT 12

05PH98_SALPA PRELIMINARY; PRT; 310 AA.

ID 05PH98_05PH98;
 DT 01-FEB-2005 (TREMBlrel. 29, Created)
 DT 01-FEB-2005 (TREMBlrel. 29, Last sequence update)
 DT 01-FEB-2005 (TREMBlrel. 29, Last annotation update)
 DE 6-phosphofructokinase isozyme.
 GN Name=pfk; OrderedlocusNames=SPA1518;
 OS Salmonella paratyphi A.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Salmonella.
 OX NCBI_Taxid=54388;
 RN [1]

NUCLEOTIDE SEQUENCE.

RC STRAIN=ATCC 9150;
 RX PubMed=15531882; DOI=10.1038/ng1470;
 RA McClelland M., Sanderson K.E., Clifton S.W., Latreille P.,
 Porwollik S., Sabo A., Meyer R., Bieri T., Ozersky P., McLellan M.,
 Harkins C.R., Wang C., Nguyen C., Bergoff A., Elliott G.,
 Kohlberg S., Strong C., Du P., Carter U., Kremizki C., Layman D.,
 Leonard S., Sun H., Fulton L., Nash W., Miner T., Mink P.,
 Delehaunty K., Fronick C., Magrini V., Nhan M., Warren W., Florea L.,
 Specht J., Wilson R.K.;
 RT "Comparison of genome degradation in Paratyphi A and Typhi, human-
 restricted serovars of Salmonella enterica that cause typhoid.";
 RL Nat. Genet. 36:1268-1274(2004).
 DR EMBL; CP000026; AAV77451.1; -; Genomic DNA.


```
QY 133 DELGGAASAFAVVASGLPPGVAADYYQORVADICRRSSTPLIDTSGGGLQHIS--GV 190
D 122 EBOYLEIESGAILVTSGLPPGVKLEKLTQLSAQAQKGRICIVDSGEALSAALAIQNI 181
QY 191 FLTKASVREIRECGSELTEPEOL-AAAHILDRGRAEVVVVSLSGGALLATRHASHR 249
D 182 ELVKPNQKELSAVLNRE-LTPDPDVRRKAAQETVNSGAKRVVSLGPGGALGVDSENCIO 240
QY 250 FSSIPMTAVSGVAGADAMVAITVGLSRGSLIKSVRLGNAGAAMLLTPGTAAACNRDV 309
D 241 VPPPVVSGSTVGAGDSMVGMATLKLAENASLEEMVRFVGAAGAAATLNGTRLCSHDDT 300
QY 310 ERF 313
D 301 QKTY 304

RESULT 14
Q83L40_SHIFL PRELIMINARY; PRT; 309 AA.
ID Q83L40_SHIFL;
AC Q83L40; Q7UC13;
DT 01-JUN-2003 (TREMBLrel. 24, last sequence update)
DT 01-JUN-2003 (TREMBLrel. 31, last annotation update)
DT 13-SEP-2005 (TREMBLrel. 31, last annotation update)
DE 6-phosphofructokinase II.
GN Name=pfkB; OrderedLocNames=81624, SFL507;
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxId=623;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=301 / Serotype 2a;
RX MEDLINE=22272406; PubMed=12384590; DOI=10.1093/nar/gkf566;
RA Jin O., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA Sun L., Xue Y., Zhao Y., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA Yu J.;
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RT through comparison with genomes of Escherichia coli K12 and O157.";
RL Nucleic Acids Res. 30:4432-4441(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=2457T / ATCC 700930 / Serotype 2a;
RX MEDLINE=22590274; PubMed=12704152;
RA DOI=10.1128/JAI.71.5.2775-2786.2003;
RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
RA Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
RA Schwartz D.C., Blattner F.R.;
RT "Complete genome sequence and comparative genomics of Shigella
RT flexneri serotype 2a strain 2457T.";
RL Infect. Immun. 71:2775-2786(2003).
DR EMBL; AB005674; AAN43098.2; -; Genomic DNA.
DR EMBL; AB016983; AAP16988.1; -; Genomic DNA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR InterPro; IPR002173; pfkB.
DR InterPro; IPR011611; pfkB_region.
DR Pfam; PF00294; pfkB; 1.
DR PROSITE; PS00583; pfkB_KINASES_1; 1.
DR PROSITE; PS00584; pfkB_KINASES_2; 1.
KM Complete proteome; Kinase.
SQ SEQUENCE 309 AA; 32456 MW; 898662D90585533 CRC64;

Query Match 30.8%; Score 530.5; DB 2; Length 309;
Best Local Similarity 41.1%; Pred. No. 3,7e-28;
Matches 125; Conservative 54; Mismatches 120; Indels 5; Gaps 4;
```

```
QY 73 GSTGSLIMALLGDAGVPPFVPIPIAASSTRESFTYNESRTAKQYFVLPGBSLTVAEDPQL 132
D 63 GATGEHLVSLIADENPVATVEAKDWTRQNIHVEVNSGEOYFVPMGALINDEFRO-L 121
QY 133 DELGGAASAFAVVASGLPPGVAADYYQORVADICRRSSTPLIDTSGGGLQHIS--GV 190
D 122 EBOYLEIESGAILVTSGLPPGVKLEKLTQLSAQAQKGRICITIDSSEALSAALAIQNI 181
QY 191 FLTKASVREIRECGSELTEPEOL-AAAHILDRGRAEVVVVSLSGGALLATRHASHR 249
D 182 ELVKPNQKELSAVLNRE-LTPDPDVRRKAAQETVNSGAKRVVSLGPGGALGVDSENCIO 240
QY 250 FSSIPMTAVSGVAGADAMVAITVGLSRGSLIKSVRLGNAGAAMLLTPGTAAACNRDV 309
D 241 VPPPVVSGSTVGAGDSMVGMATLKLAENASLEEMVRFVGAAGAAATLNGTRLCSHDDT 300
QY 310 ERF 313
D 301 QKTY 304

RESULT 15
Q8XE22_ECO57 PRELIMINARY; PRT; 309 AA.
ID Q8XE22_ECO57;
AC Q8XE22; Q7ADG0;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, last sequence update)
DT 13-SEP-2005 (TREMBLrel. 31, last annotation update)
DE 6-phosphofructokinase II; suppressor of pfkA (6-phosphofructokinase
DE II).
GN Name=pfkB; OrderedLocNames=EC2429, 22752;
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxId=83334;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=O157:H7 / EDL933 / ATCC 700927 / EHEC;
RX MEDLINE=21074935; PubMed=11206551; DOI=10.1038/35054089;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasser J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackert J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grodeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potomousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533(2001).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=O157:H7 / Sakai / RIMD 0509952 / EHEC;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Onishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
DR EMBL; AB005174; AAG56709.1; -; Genomic DNA.
DR EMBL; BA000007; BAB35852.1; -; Genomic DNA.
DR PIR; A85781; A85781.
DR PIR; E90932; E90932.
DR GO; GO:0016301; F:kinase activity; IEA.
DR InterPro; IPR002173; pfkB.
DR InterPro; IPR011611; pfkB_region.
DR Pfam; PF00294; pfkB; 1.
DR PROSITE; PS00583; pfkB_KINASES_1; 1.
DR PROSITE; PS00584; pfkB_KINASES_2; 1.
KM Complete proteome; Kinase.
SQ SEQUENCE 309 AA; 32456 MW; 898662D90585533 CRC64;

Query Match 30.8%; Score 530.5; DB 2; Length 309;
```


Best Local Similarity 41.1%; Pred. No. 3.7e-28;
Matches 125; Conservative 54; Mismatches 120; Indels 5; Gaps 4

```
QY      13 RIITLTMPALUITTSVDVVRPTEKKRCGAPRPDGGGGINVARIVHIGCCSTLPPAG 72
       |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      3 RYIYTLTAVSLDSAITTPOIYPBEGKICRTAPVEPEGGGINVARIAHLIGGSATLFPAG 62
       |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY      73 GSTGSLTMLLGDACVPFRVIPAASTRESFTVNESRTAKOYRFVLPGPELTAEOBQL 132
       |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      63 GATGEHLVSLLDENVPAATVEAKOMTRONLHVHVAASGEQIRFVMFGAALNDEFRO-L 121
       |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY      133 DELRGAAASAFAVFVASGSLPPGYAADYORVADICRRSPTLLIDTSGGGLQHISS--GV 190
       |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      122 BEQVAVEISSGALLIVISGSLPPGYKLEKLQLLSAQNQGIRCIIDBSGELSAALAIIGN 181
       |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY      191 FLTKASVRELRECVSGSEILITEPBOL-AAAHELIDRGAEVVVVSLGSCGALLATRHASHR 249
       |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      182 ELVKNDQNELSATLVNRE-LTOPDDYRKAAQEIIVNSGKARKVVVSLGPQGLGVDSENCIO 240
       |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY      250 PSSIMPTAVSGAGDAMVAALTIVGSLRGSLSIKSRLENAGAAMLLTPGTAAACNRDY 309
       |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      241 VVPPEVKSQSTVGAGDSNVGMTLKLAENASILEMVRFGVAAGSAATLTNQSTRCSHDDT 3000
       |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY      310 ERFF 313
       ::::
Db      301 QKIY 304
```

Search completed: March 23, 2006, 05:24:17
Job time : 88.7035 secs

This Page Blank (uspto)

RESULT 2
US-09-489-039A-10377
; Sequence 10377, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/489,039A
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 10377
; LENGTH: 314
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-10377

Query Match 30.3%; Score 520.5; DB 2; Length 314;
Best Local Similarity 39.6%; Pred. No. 8.2e-45;
Matches 122; Conservative 52; Mismatches 131; Indels 3; Gaps 2;

QY 9 EGRRIITLTNNPALDITTSVDVVRPTEKRCGAPRYDPGGGGINVARIYHVLGGCSTAL 68
DB 3 EGMKXITVTLTLPISDSTQTPQIYPEGKRCSPAIFVEPGGGGINVAAVTFLLGKATAI 62
QY 69 PPAAGSTGSLMALGDAVGFPRVPIPAASTRESFTVNESTAOYFVLPBPSLTVAEQ 128
DB 63 PPAAGSTGSLMALGDAVGFPRVPIPAASTRESFTVNESTAOYFVLPBPSLTVAEQ 122
QY 129 EOCDELRGAAASAFVVASGSLPPGVADYQYRVADICRRSSTPLIDTSGGGLQHI-- 186
DB 123 RR-LEEKVLTTEPSGLVYSSLPISVDLMQVKNACQGRCLIDSSGDALAAALD 181
QY 187 SSGVFLKASVRELRCEVSELLTEPEQLAAAHLEIDRGRAEVVVSIGSGALLATRHA 246
DB 182 VGNIELVXPNGKELSAVQRLSQPDVRLAAQSLIGSKVRVVSIGPGALGVDSAG 241
QY 247 SHRFSIEMTAVSGVAGDAMVAATVGLSRGMSLISVRLGNAAAMLTPTAACNR 306
DB 242 SVQVVPPEPKSQSTVAGDSVMGAMTURLAENASLEIDVVRGVAAASMTINQTRLCR 301
QY 307 DVERFE 314
DB 302 ANTKITD 309

RESULT 3
US-09-134-000C-6381
; Sequence 6381, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE REFERENCE: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/134,000C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6381
; LENGTH: 318
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-6381

Query Match 18.7%; Score 322; DB 2; Length 318;
Best Local Similarity 29.6%; Pred. No. 1.9e-24;

Matches 85; Conservative 51; Mismatches 141; Indels 10; Gaps 4;
QY 14 IITLNNPALDITTSVDVVRPTEKRCGAPRYDPGGGGINVARIYHVLGGCSTALP 73
DB 7 IVTVNMPISIDISLVDLHKLDTVNRFSQYTKTPGGKGLVTVFVHDLGADVATATGVLG 66
QY 74 STGSLMALGDAVGFPRVPIPAASTRESFTVNESTAOYFVLPBPSLTVAEQ 133
DB 67 FHGAFINELKKNIIPAFSTISKEETRDSIAI--LHEGNTTEILEGPTVSPESINPLE 124
QY 134 ELKGAASAFVVASGSLPPGVADYQYRVADICRRSSTPLIDTSGGGLQHISSG---V 190
DB 125 NPDOLIQAEIVITISGLAGLSPDFQELVQAAHQEVVVLDTGSDLSRLQVQWKP 164
QY 191 FLKASVRELRCEVSELLTEPEQLAAAHLEIDR--GRAEVVVSIGSGALLATRHA 247
DB 185 YLKPNELEELGQDFSEN--LAAVQYALTKPMPAGIEWIVISIGKGAIAKHHDQF 242
QY 248 HRFSIEMTAVSGVAGDAMVAATVGLSRGMSLISVRLGNAAAMLTPTAACNR 294
DB 243 YRVKIPITQAKNPVSGDATTIAGLADAPABAILKWMAGMA 289

RESULT 4
US-09-134-000C-3807
; Sequence 3807, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE REFERENCE: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/134,000C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3807
; LENGTH: 321
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-3807

Query Match 17.3%; Score 298; DB 2; Length 321;
Best Local Similarity 29.5%; Pred. No. 5.6e-22;
Matches 88; Conservative 59; Mismatches 139; Indels 12; Gaps 6;

QY 10 GKRPIITLTNNPALDITTSVDVVRPTEKRCGAPRYDPGGGGINVARIYHVLGGCSTALP 69
DB 14 GKQMIYTVTLNPSIDIFVHVDHLQIGDLNMTNDPFLPGGKGINVARIILKRMETESTALG 73
QY 70 PPAAGSTGSLMALGDAVGFPRVPIPAASTRESFTVNESTAOYFVLPBPSLTVAEQ 129
DB 74 FLGGFTGSPFADWLKKEELQTNFTPSADRIKIK--KSTETEINGL--GPAITVEE-- 128
QY 130 OCBDELRGAAASAFVVASGSLPPGVADYQYRVADICRRSSTPLIDTSGGGLQHI-- 185
DB 129 --IQELKQAVSRQAGDIVLSSSTASLRKGYEBELIQVKKGAEPVIDTIGEDLMA 186
QY 186 ISSGVFLKASVRELRCEVSELLTEPEQLAAAHLEIDRGRAEVVVSIGSGALLATRHA 245
DB 187 LSQKPLLVKKNHLELAEVHTFTSIEDLIPYGRILLEEG--AQHVIIISMAGDALFTTE 245
QY 246 ASHRFSIEMTAVSGVAGDAMVAATVGLSRGMSLISVRLGNAAAMLTPTAACNR 303
DB 246 GYVRSNVLKRLPLKNSVAGDSMTAGFTGNFSKQDPLEAKFWAGVAGSATATFSDILAS 303

RESULT 5
US-09-134-001C-5583
; Sequence 5583, Application US/09134001C
; Patent No. 6380370

```

; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 5583
; LENGTH: 312
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5583

Query Match      16.5%; Score 284; DB 2; Length 312;
Best Local Similarity 26.8%; Pred. No. 1.5e-20;
Matches 80; Conservative 68; Mismatches 141; Indels 10; Gaps 6;

QY 14 IITLTMNPALDITTSVDVVRPTEKMRGCGAPRYDPPGGGINVARIYHVILGCGSTALFPAGG 73
DB 8 IYTVNPSPIDYIITFDGPEISGLNRAKATYKFAAGKGINVSRVLTLDVDTALGFSGG 67

QY 74 STGSLMALIGDAGVPRFVPIAASTRESFTVN--ESRTAKQYRFVLPGPSLTVAEOEQL 132
DB 68 PFGGFIQTLEDSNIQSDPVGVEDTR---INVKLKSGGTEINAGPKYTHAQFEDLL 123

QY 133 DELFGAASAFAFVVASGSLPPGVADYQYRVADICRSTPLIIDTSGGGLQHI--SSGVF 191
DB 124 SQIR-RTTNDIVIVASGVSPNSIPSDAYAQITKGTGLVDVDAEKDLVEVLPPYRPL 182

QY 192 LKASVRELRECVSELTPEOLAANHELIDGRABRVVVSOGALLATRHASHRES 251
DB 183 FIKKQDELEWENTYVKSDEVIKIKGELIKKG--AQSVIISLGGDGLTYDDQHSIKAV 241

QY 252 SIPTAVSGVAGDAMVAATITVGLSRGMSLISYRLNAGAAAMLLTPGTAAACNRDDVE 310
DB 242 NPGSHVNVTVSGSDSTVAGVAGSLMGLINIDEAFQAVASGTATAFSEDLA--TRDAIE 298

RESULT 6
US-09-715-040-2
; Sequence 2, Application US/09715040
; Patent No. 6806068
; GENERAL INFORMATION:
; APPLICANT: Mockel, Bettina
; APPLICANT: Pfeifferle, Walter
; TITLE OF INVENTION: NEW NUCLEOTIDE SEQUENCES WHICH ENCODE THE PFK GENE
; FILE REFERENCE: 21123/275321
; CURRENT APPLICATION NUMBER: US/09/715,040
; CURRENT FILING DATE: 2001-06-04
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 330
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-715-040-2

Query Match      16.2%; Score 279; DB 2; Length 330;
Best Local Similarity 31.2%; Pred. No. 5.2e-20;
Matches 102; Conservative 46; Mismatches 133; Indels 46; Gaps 12;

QY 14 IITLTMNPALDITTSVDVVRPTEKMRGCGAPRYD-----PGGGGINVARIYHVILGCGSTAL 68
DB 2 IITFPNPSPIDSTSLG----EELSRGVSQRLDSVTAVAGGKGINVAHAVALAFETILAV 57

QY 69 FPAAGSTGSLMALIGDAGVPRFVPIAASTRESFTVNE--SRTAKQYRFVLPGPSLTVA 126
DB 58 FPAAGL--DPFVPLVRDIGLPEVETVINKVNTITVTEPDGTTK-----LNGPAPLS 110

; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
```

```

QY 127 EQ-----EQCLDELRGAAASAFVVASGSLPPGVADYQYRVADICR--RSSTPLILDT 178
DB 111 EQKRLSLEKVLIDALR---PEVTWVVLGSLPFGAPVDWYARLTALIHSAKRPDVRVAVDT 167

QY 179 S-----GGGLQHSISGVLLKASVRELRECV---GSELTTEPEQ-----LAAAHLEI 222
DB 168 SDRPLMALGESLDTTGGAPNLIKPNGLQLANTDGEELBARAAGDYDAI1AAADVLV 227

QY 223 DRGRAEVVVVSLGSGALLATRHASHRESFSIPTAVSGVAGDAMVAATITVGLSRGMSLI 282
DB 228 NRG-IEQVLTVLGAAGAVLVNAGCAWTATSPKIDIVSTVAGDCCALGFWMASSQKKTLE 286

QY 283 KSVRLNAGAAAMLLTPGTAAACNRDDV 309
DB 287 ESLLNAYSGSTAAASLPGTTIPRPDQL 313

RESULT 7
US-10-098-626-4
; Sequence 4, Application US/10098626
; Patent No. 6921651
; GENERAL INFORMATION:
; APPLICANT: Patrick, Mike, et al.
; TITLE OF INVENTION: PROCESS FOR THE PREPARATION OF L-AMINO ACIDS BY USING CORYNEFORM
; FILE REFERENCE: 032301 WN 269
; CURRENT APPLICATION NUMBER: US/10/098,626
; CURRENT FILING DATE: 2002-03-18
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 330
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-10-098-626-4

Query Match      16.2%; Score 279; DB 2; Length 330;
Best Local Similarity 31.2%; Pred. No. 5.2e-20;
Matches 102; Conservative 46; Mismatches 133; Indels 46; Gaps 12;

QY 14 IITLTMNPALDITTSVDVVRPTEKMRGCGAPRYD-----PGGGGINVARIYHVILGCGSTAL 68
DB 2 IITFPNPSPIDSTSLG----EELSRGVSQRLDSVTAVAGGKGINVAHAVALAFETILAV 57

QY 69 FPAAGSTGSLMALIGDAGVPRFVPIAASTRESFTVNE--SRTAKQYRFVLPGPSLTVA 126
DB 58 FPAAGL--DPFVPLVRDIGLPEVETVINKVNTITVTEPDGTTK-----LNGPAPLS 110

QY 127 EQ-----EQCLDELRGAAASAFVVASGSLPPGVADYQYRVADICR--RSSTPLILDT 178
DB 111 EQKRLSLEKVLIDALR---PEVTWVVLGSLPFGAPVDWYARLTALIHSAKRPDVRVAVDT 167

QY 179 S-----GGGLQHSISGVLLKASVRELRECV---GSELTTEPEQ-----LAAAHLEI 222
DB 168 SDRPLMALGESLDTTGGAPNLIKPNGLQLANTDGEELBARAAGDYDAI1AAADVLV 227

QY 223 DRGRAEVVVVSLGSGALLATRHASHRESFSIPTAVSGVAGDAMVAATITVGLSRGMSLI 282
DB 228 NRG-IEQVLTVLGAAGAVLVNAGCAWTATSPKIDIVSTVAGDCCALGFWMASSQKKTLE 286

QY 283 KSVRLNAGAAAMLLTPGTAAACNRDDV 309
DB 287 ESLLNAYSGSTAAASLPGTTIPRPDQL 313

; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
```


Db 64 GGALQYIEBQIEETTRIKQAFKIKGETRNCIAL-LHEGOQTEILEKGPTEILEKESEEF 121
Qy 132 LDELRCGAASAFAVVASGSLPPGVADYYQVRADICRRSSTPLILDTSGGGLQHI-SSGV 190
Db 122 KSHLLKPKETNDVAVMSSGLPKGINTDYADIVLAKEGGLITLIDSSGSGLEEVLSNV 181
Qy 191 --FLIKASVRELRECVSGSELLTEPBQLAAHELIDRGRAEVVVSLSGQALLATRHASH 248
Db 182 KPTVYKPNIDELSQLINKYKTVNDIKELKAAVSOPIFNDIEMIIVSLGSEGAFAKHQKXY 241
Qy 249 RFSISPMTAVSGVAGDAMVAITVGLSRGSLIKSVLGNAGAAAMLITGTAACRRD 308
Db 242 KVNIPNKKVNPVSGSDSTVAGIASGLIHQOTDELBKKAFAFMLNAMEOQCHINTDK 301
Qy 309 VERFFE 314
Db 302 FDEIFK 307
RESULT 11
US-09-107-433-4601
; Sequence 4601, Application US/09107433
; Patent No. 6800744
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
; SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE FOR DIAGNOSTIC
; THERAPEUTICS
; NUMBER OF SEQUENCES: 5206
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: <Unknown>
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: <Unknown>
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107.433
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/ 085131
; FILING DATE: May 12, 1998
; APPLICATION NUMBER: 60/051553
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Arinello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 4601:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 310 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULAR TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...310
; SEQUENCE DESCRIPTION: SEQ ID NO: 4601:
US-09-107-433-4601
Query Match 15.1%, Score 260, DB 2, Length 310,
Best Local Similarity 28.0%, Pred. No. 4,2e-18,

Matches 81; Conservative 51; Mismatches 137; Indels 20; Gaps 7;
Qy 12 PRITLTNMPALDITTSVDVVRPTKRRCGAPRYDPGGGGINVARIYHVLGGCSTALFPA 71
Db 1 PMILTVMNNSIDISYLDLDELKIDTVNRVVDVTKTAGKGLNTRVLRSEGDSTLARGLV 60
Qy 72 GGSSTGLMLALLDAGYPPFVVIPIASTRESFTV---NSRTRAKYRFVLPQSLTVAE 127
Db 61 GSKGEFLVHIDQ--VKDQFSTIOGETRNCIALHMDNQTVELEK-----GEVLEOE 113
Qy 128 QEGCLDELRCGAASAFAVVASGSLPPGVADYYQVRADICRRSSTPLILDTSGGGLQHI 187
Db 114 GQDFLEHFKCLLESVEVAISGLPAGLPDYTASLVELANQAKLVVLDSCGALQAVL 173
Qy 188 SGVF---LTKASVRELRECVSGSELLTEPBQL-AAHELIDRGRAEVVVSLSGQALLAT 243
Db 174 ESPHKPTVYKPNNEELSQLGREVSEDELEKVELQSLFTG-IEWIIVSLGANGTAKH 232
Qy 244 RHASHRFSISPMTAVSGVAGDAMVAITVGL-----SRGSLIKSVRLG 288
Db 233 GDTFFYKVDIPRIQVNPVSGSDSTVAGISSGLHKESDAGLLIKANVLG 281
RESULT 12
US-08-961-539-2
; Sequence 2, Application US/08961539
; Patent No. 5861281
; GENERAL INFORMATION:
; APPLICANT: Zalacain, Magdalena
; APPLICANT: Brown, James R.
; TITLE OF INVENTION: No. 5861281el 1acc
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dechert, Price & Rhoads
; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103-2793
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,539
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Dickinson, Q. Todd
; REGISTRATION NUMBER: 28,354
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215/994-2252
; TELEFAX: 215/994-2222
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 303 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-961-539-2
Query Match 14.0%, Score 240, DB 1, Length 303,
Best Local Similarity 27.4%, Pred. No. 4,5e-16;
Matches 78; Conservative 54; Mismatches 141; Indels 12; Gaps 5;
Qy 14 ITLTNMPALDITTSVDVVRPTKRRCGAPRYDPGGGGINVARIYHVLGGCSTALFPA 73

Db 2 IYVTLNPSIDYIRLQOVKGVNRMDSDDKFAAGCKINVRVLRKLNISNTATGFIIG 61
Qy 74 STGSLMALIGDAGVPRVPIPIASTRESFTVNESRTAKQYRFLPGPSLTVAEQECLO 133
Db 62 FTGKFTIDTLAEERIEFRFQVADTRINVKIKADQETE-----INGTGPVEPVK--LE 114
Qy 134 ELRGAAASAA--FVVASGSLPPGVADYQYRVADICRRSSTPLIDTSGGGL-QHISG 189
Db 115 ELKAILSSLTAEDTVVFAAGSSAKNLGNVIYKDLISLRTQGAQVCDPFEQOTLIDSLDYQ 174
Qy 190 VFLKASVRELRECVSESLTEPEQLAAHELDGRARVVVYVSLSGQALLATRRASHR 249
Db 175 PLLVKPNNHGALFGVKLSLDEIEKYARELLAKG-AQNVIIISMAGDGLVTSBGAYF 233
Qy 250 FSSIPMTAVSGVAGDAMVAITVGLSRGWSLIKSVRLGNAGAA 294
Db 234 AKPIKGTIVKNSVGAQDSMVAGFTGEFVKSQDAVEAFKMGVACGTA 278

RESULT 13
US-09-185-826-2
; Sequence 2, Application US/09185826
; Patent No. 6171840
; GENERAL INFORMATION:
; APPLICANT: Zalecain, Magdalena
; TITLE OF INVENTION: No. 6171840e1 1acc
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dechert, Price & Rhoads
; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103-2793
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/185.826
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/961.539
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Dickinson, Q. Todd
; REGISTRATION NUMBER: 28,354
; REFERENCE/DOCKET NUMBER: GM10114
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215/994-2252
; TELEFAX: 215/994-2222
; TEXT:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 303 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-185-826-2

Query Match 14.0%; Score 240; DB 2; Length 303;
Best local Similarity 27.4%; Pred. No. 4.5e-16;
Matches 78; Conservative 54; Mismatches 141; Indels 12; Gaps 5;
Qy 14 IITLNNPALDITTSVDVVRPTEKRCGAPRYDPGGGGINVARIHVHVGSCSTALFPAGG 73
Db 2 IYVTLNPSIDYIRLQOVKGVNRMDSDDKFAAGCKINVRVLRKLNISNTATGFIIG 61
Qy 74 STGSLMALIGDAGVPRVPIPIASTRESFTVNESRTAKQYRFLPGPSLTVAEQECLO 133

Db 62 FTGKFTIDTLAEERIEFRFQVADTRINVKIKADQETE-----INGTGPVEPVK--LE 114
Qy 134 ELRGAAASAA--FVVASGSLPPGVADYQYRVADICRRSSTPLIDTSGGGL-QHISG 189
Db 115 ELKAILSSLTAEDTVVFAAGSSAKNLGNVIYKDLISLRTQGAQVCDPFEQOTLIDSLDYQ 174
Qy 190 VFLKASVRELRECVSESLTEPEQLAAHELDGRARVVVYVSLSGQALLATRRASHR 249
Db 175 PLLVKPNNHGALFGVKLSLDEIEKYARELLAKG-AQNVIIISMAGDGLVTSBGAYF 233
Qy 250 FSSIPMTAVSGVAGDAMVAITVGLSRGWSLIKSVRLGNAGAA 294
Db 234 AKPIKGTIVKNSVGAQDSMVAGFTGEFVKSQDAVEAFKMGVACGTA 278

RESULT 14
US-09-583-110-5072
; Sequence 5072, Application US/09583110
; Patent No. 6699703
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; TITLE OF INVENTION: Pneumoniae for Diagnostic and Therapeutics
; FILE REFERENCE: PAT00-07A
; CURRENT APPLICATION NUMBER: US/09/583.110
; CURRENT FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107.433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085.131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051.553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 5072
; LENGTH: 303
; TYPE: PRN
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (17)
; OTHER INFORMATION: Identity of amino acid sequences at the above locations are unknown
US-09-583-110-5072

Query Match 13.4%; Score 230; DB 2; Length 303;
Best local Similarity 27.0%; Pred. No. 4.8e-15;
Matches 77; Conservative 53; Mismatches 143; Indels 12; Gaps 5;
Qy 14 IITLNNPALDITTSVDVVRPTEKRCGAPRYDPGGGGINVARIHVHVGSCSTALFPAGG 73
Db 2 IYVTLNPSIDYIRLQOVKGVNRMDSDDKFAAGCKINVRVLRKLNISNTATGFIIG 61
Qy 74 STGSLMALIGDAGVPRVPIPIASTRESFTVNESRTAKQYRFLPGPSLTVAEQECLO 133
Db 62 FTGKFTIDTLAEERIEFRFQVADTRINVKIKADQETE-----INGTGPVEPVK--LE 114
Qy 134 ELRGAAASAA--FVVASGSLPPGVADYQYRVADICRRSSTPLIDTSGGGL-QHISG 189
Db 115 ELKAILSSLTAEDTVVFAAGSSAKNLGNVIYKDLISLRTQGAQVCDPFEQOTLIDSLDYQ 174
Qy 190 VFLKASVRELRECVSESLTEPEQLAAHELDGRARVVVYVSLSGQALLATRRASHR 249
Db 175 PLLVKPNNHGALFGVKLSLDEIEKYARELLAKG-AQNVIIISMAGDGLVTSBGAYF 233
Qy 250 FSSIPMTAVSGVAGDAMVAITVGLSRGWSLIKSVRLGNAGAA 294
Db 234 AKPIKGTIVKNSVGAQDSMVAGFTGEFVKSQDAVEAFKMGVACGTA 278

Job time : 24.2209 secs

APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE FOR DIAGNOSTIC
THERAPEUTICS

NUMBER OF SEQUENCES: 5206

CORRESPONDENCE ADDRESS:

ADDRESSEE: GENOME THERAPEUTICS CORPORATION

STREET: 100 Beaver Street

CITY: Waltham

STATE: Massachusetts

COUNTRY: USA

ZIP: 02354

COMPUTER READABLE FORM:

MEDIUM TYPE: CD-ROM ISO9660

COMPUTER: <Unknown>

OPERATING SYSTEM: <Unknown>

SOFTWARE: <Unknown>

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/107,433

FILING DATE: 30-Jun-1998

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/085131

FILING DATE: May 12, 1998

APPLICATION NUMBER: 60/051553

FILING DATE: July 2, 1997

ATTORNEY/AGENT INFORMATION:

NAME: Arinello, Pamela Deneke

REGISTRATION NUMBER: 40,489

REFERENCE/DOCKET NUMBER: GTC-011

TELECOMMUNICATION INFORMATION:

TELEPHONE: (781)893-8277

TELEFAX: (781)893-5007

INFORMATION FOR SEQ ID NO: 4423:

SEQUENCE CHARACTERISTICS:

LENGTH: 311 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: YES

ORIGINAL SOURCE:

ORGANISM: Streptococcus pneumoniae

FEATURE:

NAME/KEY: misc feature

LOCATION: (B) LOCATION 1...311

SEQUENCE DESCRIPTION: SEQ ID NO: 4423:

US-09-107-433-4423

Query Match 13.4%; Score 230; DB 2; length 311;
Best Local Similarity 27.0%; Pred. No. 5e-15;
Matches 77; Conservative 53; Mismatches 143; Indels 12; Gaps 5;

14 IITITNPADDITTSVVVPRTEKRCGAPRYPDGGGINVARIYHVLGGSTALFPAGG 73
10 IYVTLNPISIDYIRXQVKGVSVNRMDSDKDFAGEKGINSVRLKRLNLSMTATGFTGG 69
74 STGSLMALGDAVPPRVPPIASTRESFTVNSRTAKQRFVLPGPSLVAEQEQLD 133
70 FTGKFTIDTLAEEIETRFVQVAEDTRINVKIKADQETE----INGTGPTEPVK--LE 122
134 ELRGAASAA--FVVASGSLPGVAADYQYRVADICRSSTPLILDTSGGGL-OHISG 189
123 ELKAILSLTAEDTTFVAGSSAKNLGNVYKDLISLRTQGAQVCDPEGQTLIDSLDQ 182
190 VFLLKASVREIRECVSEELTEPEQLAAHMLIDRGRAEVVVVLSGQALATRAHSHR 249
183 PLVTKPNNHLEIGALFGVLESLDEIEKVARBLAKG-AQNVYISMADGALLVTSBGAYF 241
250 FSSIPMTAVSGAGADVAITVGLSGNGLIKSVRLGNAGAA 294
242 AKPIKGTIVKSVGAGDSNVAGFTGSEFVKSQDAVAFKMGVACGTA 286

Search completed: March 23, 2006, 05:30:10

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: March 23, 2006, 06:14:33 ; Search time 71.2896 Seconds
(without alignment)
1986.885 Million cell updates/sec

Title: US-10-617-038-22

Perfect score: 1720
Sequence: 1 MTEPAAWDEKPRITLITM.....TEVGQDYVHPIVNPASP 339

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA_Main:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
- 3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
- 4: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
- 6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1720	100.0	339	4	US-10-282-122A-62568
2	1720	100.0	339	4	US-10-282-122A-64633
3	1720	100.0	339	4	US-10-617-038-22
4	726	42.2	310	4	US-10-369-493-19155
5	617.5	35.9	308	4	US-10-369-493-10209
6	587.5	34.7	505	4	US-10-282-122A-50578
7	587.5	34.2	338	4	US-10-282-122A-48067
8	534.5	31.1	310	3	US-09-815-242-13976
9	531.5	30.9	309	3	US-09-815-242-10180
10	531.5	30.9	309	4	US-10-369-493-851
11	531.5	30.9	309	4	US-10-282-122A-55670
12	531.5	30.9	311	4	US-10-282-122A-55963
13	355	20.6	317	4	US-10-156-761-11226
14	330	19.2	309	4	US-10-156-761-11476
15	311	18.1	303	4	US-10-282-122A-45630
16	304	17.7	309	4	US-10-282-122A-52089
17	302	17.6	1249	5	US-10-450-763-52025
18	294	17.1	304	5	US-10-501-282-262
19	290	16.9	305	3	US-09-815-242-10787
20	289	16.8	304	5	US-10-873-467-74
21	289	16.8	305	4	US-10-282-122A-70669
22	284	16.5	306	4	US-10-282-122A-70667
23	284	16.5	312	4	US-10-724-972A-7339
24	283	16.5	306	4	US-10-282-122A-44238
25	282	16.4	303	3	US-09-815-242-5393
26	282	16.4	306	3	US-09-815-242-12300
27	279.5	16.2	312	5	US-10-501-282-1494

28	279	16.2	330	3	US-09-738-626-5624	Sequence 5624, Ap
29	279	16.2	330	4	US-10-098-626-4	Sequence 4, Appl1
30	279	16.2	330	4	US-10-781-014-58	Sequence 58, Appl
31	279	16.2	330	5	US-10-494-836-58	Sequence 58, Appl
32	276.5	16.1	302	3	US-09-815-242-4955	Sequence 4955, Ap
33	276.5	16.1	308	3	US-09-815-242-10775	Sequence 10775, A
34	276.5	16.1	308	4	US-10-282-122A-42555	Sequence 42555, A
35	272	15.8	304	4	US-10-282-122A-57485	Sequence 57485, A
36	270	15.7	303	4	US-10-282-122A-45645	Sequence 45645, A
37	269.5	15.7	304	4	US-10-282-122A-59615	Sequence 59615, A
38	266	15.5	306	4	US-10-369-493-17156	Sequence 17156, A
39	266	15.5	307	4	US-10-282-122A-60536	Sequence 60536, A
40	264.5	15.4	314	4	US-10-724-972A-6354	Sequence 6354, Ap
41	264	15.3	306	4	US-10-282-122A-71925	Sequence 71925, A
42	263.5	15.3	307	4	US-10-282-122A-47216	Sequence 47216, A
43	260	15.1	310	5	US-10-617-320-4601	Sequence 4601, Ap
44	252.5	14.7	321	4	US-10-282-122A-53758	Sequence 53758, A
45	250	14.5	309	5	US-10-472-928-2382	Sequence 2382, Ap

ALIGNMENTS

RESULT 1
US-10-282-122A-62568
Sequence 62568, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zykkind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: EUTRA-034A
CURRENT FILING DATE: 2003-02-20
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 62568
LENGTH: 339
TYPE: PRT
ORGANISM: Mycobacterium bovis
US-10-282-122A-62568

```
Query Match      100.0%; Score 1720; DB 4; Length 339;
Best Local Similarity 100.0%; Pred. No. 7.1e-154;
Matches 339; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTEPAAWDEGKPRITLTMMNPAIDITTSVDVVRPTEKMRGAGARYDPGGGGINVARIYHV 60
DB 1 MTEPAAWDEGKPRITLTMMNPAIDITTSVDVVRPTEKMRGAGARYDPGGGGINVARIYHV 60
QY 61 LGGCSTALFPAGSGTGSILMALLDGAGVPRFVPIPIASTRESFTVNESRTAKQYRFLVPG 120
DB 61 LGGCSTALFPAGSGTGSILMALLDGAGVPRFVPIPIASTRESFTVNESRTAKQYRFLVPG 120
QY 121 PSLTVAEQECCLDELRGAAASAAFFVVASGSLPGVADYQYVADICRRSSTPLILDTSG 180
DB 121 PSLTVAEQECCLDELRGAAASAAFFVVASGSLPGVADYQYVADICRRSSTPLILDTSG 180
QY 181 GGLQHISSGVFLKASVREIRECVGSELTPEQOLAAHEILDRGRAEVVVVSLGSGCAL 240
DB 181 GGLQHISSGVFLKASVREIRECVGSELTPEQOLAAHEILDRGRAEVVVVSLGSGCAL 240
QY 241 LATRHASHRFSIIPMTAVSGVAGDAMVAAITVGLSRGWSLKSIVRLGNAGAAMLTPG 300
DB 241 LATRHASHRFSIIPMTAVSGVAGDAMVAAITVGLSRGWSLKSIVRLGNAGAAMLTPG 300
QY 301 TAAACNRDDVERFFELAAPEPTVGODQYVWHPINPEASP 339
DB 301 TAAACNRDDVERFFELAAPEPTVGODQYVWHPINPEASP 339

RESULT 2
US-10-282-122A-64633
; Sequence 64633, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zykend, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forevych, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
```

```
; SEQ ID NO 64633
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-282-122A-64633
Query Match      100.0%; Score 1720; DB 4; Length 339;
Best Local Similarity 100.0%; Pred. No. 7.1e-154;
Matches 339; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTEPAAWDEGKPRITLTMMNPAIDITTSVDVVRPTEKMRGAGARYDPGGGGINVARIYHV 60
DB 1 MTEPAAWDEGKPRITLTMMNPAIDITTSVDVVRPTEKMRGAGARYDPGGGGINVARIYHV 60
QY 61 LGGCSTALFPAGSGTGSILMALLDGAGVPRFVPIPIASTRESFTVNESRTAKQYRFLVPG 120
DB 61 LGGCSTALFPAGSGTGSILMALLDGAGVPRFVPIPIASTRESFTVNESRTAKQYRFLVPG 120
QY 121 PSLTVAEQECCLDELRGAAASAAFFVVASGSLPGVADYQYVADICRRSSTPLILDTSG 180
DB 121 PSLTVAEQECCLDELRGAAASAAFFVVASGSLPGVADYQYVADICRRSSTPLILDTSG 180
QY 181 GGLQHISSGVFLKASVREIRECVGSELTPEQOLAAHEILDRGRAEVVVVSLGSGCAL 240
DB 181 GGLQHISSGVFLKASVREIRECVGSELTPEQOLAAHEILDRGRAEVVVVSLGSGCAL 240
QY 241 LATRHASHRFSIIPMTAVSGVAGDAMVAAITVGLSRGWSLKSIVRLGNAGAAMLTPG 300
DB 241 LATRHASHRFSIIPMTAVSGVAGDAMVAAITVGLSRGWSLKSIVRLGNAGAAMLTPG 300
QY 301 TAAACNRDDVERFFELAAPEPTVGODQYVWHPINPEASP 339
DB 301 TAAACNRDDVERFFELAAPEPTVGODQYVWHPINPEASP 339

RESULT 3
US-10-617-038-22
; Sequence 22, Application US/10617038
; Publication No. US20040057963A1
; GENERAL INFORMATION:
; APPLICANT: Andersen, Peter
; APPLICANT: Rosenkrands, Ida
; APPLICANT: Scrym, Anette
; TITLE OF INVENTION: Therapeutic TB Vaccine
; FILE REFERENCE: SSISAUSA
; CURRENT APPLICATION NUMBER: US/10/617,038
; CURRENT FILING DATE: 2003-07-11
; PRIOR APPLICATION NUMBER: DK PA 2002 01098
; PRIOR FILING DATE: 2002-07-13
; PRIOR APPLICATION NUMBER: US 60/401,725
; PRIOR FILING DATE: 2002-08-07
; NUMBER OF SEQ ID NOS: 187
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 22
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-617-038-22
Query Match      100.0%; Score 1720; DB 4; Length 339;
Best Local Similarity 100.0%; Pred. No. 7.1e-154;
Matches 339; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTEPAAWDEGKPRITLTMMNPAIDITTSVDVVRPTEKMRGAGARYDPGGGGINVARIYHV 60
DB 1 MTEPAAWDEGKPRITLTMMNPAIDITTSVDVVRPTEKMRGAGARYDPGGGGINVARIYHV 60
QY 61 LGGCSTALFPAGSGTGSILMALLDGAGVPRFVPIPIASTRESFTVNESRTAKQYRFLVPG 120
DB 61 LGGCSTALFPAGSGTGSILMALLDGAGVPRFVPIPIASTRESFTVNESRTAKQYRFLVPG 120
QY 121 PSLTVAEQECCLDELRGAAASAAFFVVASGSLPGVADYQYVADICRRSSTPLILDTSG 180
DB 121 PSLTVAEQECCLDELRGAAASAAFFVVASGSLPGVADYQYVADICRRSSTPLILDTSG 180
```

Db 121 PSLTVAEOEQCLDELRGAAASAFVVASGSLPPGVADYORVADICRRSSTPLILDTSG 180
Qy 181 GGLOHISGSGVFLKASVRELCVSGSELLTEPEQLAAHELIIDRGAEVVVSLGSGAL 240
Db 181 GGLOHISGSGVFLKASVRELCVSGSELLTEPEQLAAHELIIDRGAEVVVSLGSGAL 240
Qy 241 LATRHASHRFSIPMTAVSGVAGADAMVAITVGLSRGWSLIKSVRLGNAAGAMLTPG 300
Db 241 LATRHASHRFSIPMTAVSGVAGADAMVAITVGLSRGWSLIKSVRLGNAAGAMLTPG 300
Qy 301 TAACNRDVERFELAAPETEVGDDQYWMHPVNEAPSP 339
Db 301 TAACNRDVERFELAAPETEVGDDQYWMHPVNEAPSP 339

RESULT 4

US-10-369-493-19155
; Sequence 19155, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT FILING DATE: 2003-02-28
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 19155
; LENGTH: 310
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-10-369-493-19155

Query Match 42.2%; Score 726; DB 4; Length 310;
Best Local Similarity 49.0%; Pred. No. 7,1e-60;
Matches 152; Conservative 56; Mismatches 100; Indels 2; Gaps 2;

Qy 12 PRITLTNNPALDITTSVDVVRPTKRCGAPRYDPGGGGINVARIYVHVLGGCSTALPFA 71
Db 1 PHITLTNNPALDITTSVDVVRPTKRCGAPRYDPGGGGINVARIYVHVLGGCSTALPFA 71
Qy 72 GGSTGSLMALLDGAGVPRVPIAASSTRESFTVNESRTAKQYRFVLPGPSLTVAEOEQ 131
Db 61 GGAIGSQLELLEKRGRLRSGVPLQDARTRESFTVAEGTGAEYRFLPGPILSRERQRC 120
Qy 132 LDELRGAAASAFVVASGSLPPGVADYORVADICRRSSTPLILDTSGGGLQ-HISSGV 190
Db 121 LDALDEDAVVGATYIVASGSLAPGVPEDFYARVARGKLAIVRVVVDTPGLPRAALESGV 180
Qy 191 FLTASVRELCVSGSELLTEPEQLAAHELIIDRGAEVVVSLGSGALLATRHASHR 250
Db 181 FMAAPNRRELKDLGASVEDPATQAAARBELVAQGRAEVLVSHGADALVTTTQAQFRA 240
Qy 251 SSIPMTAVSGVAGADAMVAITVGLSRGWSLIKSVRLGNAAGAMLTPGTAACNRDDVE 310
Db 241 TPPVTRTVSSYGAGDSFVAGMTLALARGOSTEDALRWGIAAGTALTLGQIDLCYRADVE 300
Qy 311 RFF-ELAAEP 319
Db 301 RLFTQVKAEP 310

RESULT 5
US-10-369-493-10209
; Sequence 10209, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:

; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT FILING DATE: 2003-02-28
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 10209
; LENGTH: 308
; TYPE: PRT
; ORGANISM: Cytophaga hutchinsonii
US-10-369-493-10209

Query Match 35.9%; Score 617.5; DB 4; Length 308;
Best Local Similarity 43.0%; Pred. No. 1,3e-49;
Matches 131; Conservative 55; Mismatches 118; Indels 1; Gaps 1;

Qy 11 KPRITLTNNPALDITTSVDVVRPTKRCGAPRYDPGGGGINVARIYVHVLGGCSTALPFA 70
Db 2 KSFITLTLPAPDKSTSVESVQPEHKLRCDSPIYGGGGINVRAIKKLNGSSVALYP 61
Qy 71 AGSTGSLMALLDGAGVPRVPIAASSTRESFTVNESRTAKQYRFVLPGPSLTVAEOEQ 130
Db 62 IGGDTGNHFRALTEBIEQITFOIKWTRNENFVRSSTNNDFRFPMPSESEYKESWEQ 121
Qy 131 CLDELRGAAASAFVVASGSLPPGVADYORVADICRRSSTPLILDTSGGGLQ-HISSG 189
Db 122 MLAVIQPSNDFEFVVGSGTPKGVPEFTRLSQIILQKARLILDTSGALKHCKLEG 181
Qy 190 VFLKASVRELCVSGSELLTEPEQLAAHELIIDRGAEVVVSLGSGALLATRHASHR 249
Db 182 IFLCKPINEISEIVKELITKEQOEAAAMEIINSKVEILVVSIGGAGAFASKQGIYH 241
Qy 250 FSSIPMTAVSGVAGADAMVAITVGLSRGWSLIKSVRLGNAAGAMLTPGTAACNRDDV 309
Db 242 VTAPPEKRSTVAGDSMVAGMVLISRGHVEDVIKFGVACGTAATMNFETELCKLEDV 301
Qy 310 ERFFE 314
Db 302 ERLYK 306

RESULT 6
US-10-282-122A-50578
; Sequence 50578, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Maloney, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Onisen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT FILING DATE: 2003-02-20
; PRIOR FILING DATE: 2000-03-21
; PRIOR FILING DATE: 2000-03-21
; PRIOR FILING DATE: 2000-05-23

```

; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See file Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 50578
; LENGTH: 505
; TYPE: PR1
; ORGANISM: Burkholderia mallei
US-10-282-122A-50578
```

```

Query Match      34.7%; Score 597.5; DB 4; Length 505;
Best Local Similarity 41.9%; Pred. No. 2e-47;
Matches 127; Conservative 59; Mismatches 116; Indels 1; Gaps 1;
```

```

QY 12 PRITLTNNPALDITTSVDVVRPTEKMRGAPRYDPPGGGINVARIYHVLGGCSTALFPA 71
D 194 PEITLTNPAPIDAVTCEVETTRTKLRGCPARRDPGGGINVARIYHVLGGCSTALFPA 253
QY 72 GGSSTGLMALLGDAGVPRVPIPIASTRESFTVNESRTAKOYRFVLPGBSLTVABEQEC 131
D 254 GCGGTGLALRGLADEGRAGHIDLAGETRENFSLERSTGBEPFVLPGRLLAHENPRC 313
QY 132 LDELKGAASAFAFVAGSLPPGVADYYQYVADICRRSSPLILDTSGGGL-HISSGV 190
D 314 VEAIGRLADASRYLVMGSLPPGMPDDCYARLARASARGVTVVDTSGPALAALDAGV 373
QY 191 FLKASVRELRCVGSSELTEBEQLAAHSLIDGRAEVWVSLGSGALLATRHASHRF 250
D 374 YLVPSIGELRALTGLPLEDDGARLAARAVAGGRQIYALTLTGDAALVSSDDVRL 433
QY 251 SSIPMTAVSGVAGDAMVAITVGLSRGWSLKSIVRLGNAGAAMLTPGTACNRDVE 310
D 434 PGVAVAVRSALIGAGDSFVAGLVVALNNGANVADAARHALLAASASLSTGALGTRKDIA 493
QY 311 RPF 313
D 494 RIV 496

RESULT 7
US-10-282-122A-48067
; Sequence 48067, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: EUTTRA.034A
```

```

; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See file Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 48067
; LENGTH: 338
; TYPE: PR1
; ORGANISM: Burkholderia cepacia
US-10-282-122A-48067
```

```

Query Match      34.2%; Score 587.5; DB 4; Length 338;
Best Local Similarity 43.0%; Pred. No. 1e-46;
Matches 132; Conservative 48; Mismatches 124; Indels 3; Gaps 3;
```

```

QY 14 IITLTNNPALDITTSVDVVRPTEKMRGAPRYDPPGGGINVARIYHVLGGCSTALFPA 73
D 31 IYVTLNPADVATSEHIVYDTHKLCARRRDPGGGINVARTVIRLGDDCAALVLAGG 90
QY 74 STGSLMALLGDAGVPRVPIPIASTRESFTVNESRTAKOYRFVLPGBSLTVABEQEC 133
D 91 PTGDDVLTALLEAHPLSBRIRIAGETRENVCTETATGEBYRPLVMTGEMPRACAA 150
QY 134 ELKGAASAFAFVAGSLPPGVADYYQYVADICRRSSPLILDTSGGGL-HISSGV 192
D 151 RTDAMPSPRYLTLSSLPGPAPDDLYATLARTAKARGRVVDDAAGRLQALKGVL 210
QY 193 LKASVRELRCVGSSELTEBEQLAAHSLIDGRAEVWVSLGSGALLATRHASHRF 252
D 211 VKPSIGLSLALG-BPLDDTSACRKASSELVARKQADIVALTIGARKAVVTRDTRLRF 269
QY 253 IPMTAVSGVAGDAMVAITVGLSRGWSLKSIVRLGNAGAAMLTPGTACNRDVE 312
D 270 RPAVAVSTVAGDSFVGVWVWALACGVPPDDACRYVALLAASASVERGTALCTRDVERI 329
QY 313 F-ELAAE 318
D 330 HAEVLAQ 336

RESULT 8
US-09-815-242-13976
; Sequence 13976, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
```

```

; TITLE OF INVENTION: Identification of Essential Genes in
; FILE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13976
; LENGTH: 310
; TYPE: PRT
; ORGANISM: Salmonella typhi
US-09-815-242-13976

```

Query Match 31.1%; Score 534.5; DB 3; Length 310;

Best Local Similarity 41.4%; Pred. No. 9.3e-42;
Matches 126; Conservative 54; Mismatches 119; Indels 5; Gaps 4;

```

QY 13 RIITLTNNPALDITTSVDVVRPTEKRCGAPRDYGGGINVARIHVHVGCCSTALFPAG 72
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 3 RIYTLTAPSDSATTITPQIYRFBGKLRCSAVFRRGGGINVARAHLAGSATATATFPAG 62
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 73 GSTSLMALALGDAGVPRVPIPIASTRESFTVNESRTAKOYRFLPSPITVAEBOCCL 132
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 63 GATGEHVLALLADEVNPVSTVDADKQWTRQNLHVHVESGEGYRFVMPGATLDDDEFRRQ-L 121
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 133 DELGGAASAFAVVAAGSLPPGVADYQVRVADICRRSSTPLIIDTSGGGLQHTSS--GV 190
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 122 BEQVLEIESGAILVIVSGSLPPGVKVEKLTQLISAQKQICITIDSSDALTALALGDI 181
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 191 FLKASVRELRECEVSSLLTEPEQL-AAAHFLIDRGRAEVVVSIGSGALLATRHASHR 249
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 182 ELVVPNKELKLSALVNRD-LTQPDVDRKAQSLVSGKARRVVSIGPGALGISENCIQ 240
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 250 FSSIPMTAVSGVAGDANVAITVGLSRGWSLISVRLGNAAGAMLLTPTAACNRDV 309
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 241 VVPPVKSQSTVGAGDSWGMATLKLADQDASLEWVRFGVAAGSAATLNGTRLCSHDT 300
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 310 ERPF 313
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 301 QKIT 304
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

```

RESULT 9
US-09-815-242-10180

```

; Sequence 10180, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyekind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242

```

```

; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10180
; LENGTH: 309
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-815-242-10180

```

Query Match 30.9%; Score 531.5; DB 3; Length 309;

Best Local Similarity 41.1%; Pred. No. 1.6e-41;
Matches 125; Conservative 55; Mismatches 119; Indels 5; Gaps 4;

```

QY 13 RIITLTNNPALDITTSVDVVRPTEKRCGAPRDYGGGINVARIHVHVGCCSTALFPAG 72
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 3 RIYTLTAPSDSATTITPQIYRFBGKLRCSAVFRRGGGINVARAHLAGSATATATFPAG 62
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 73 GSTSLMALALGDAGVPRVPIPIASTRESFTVNESRTAKOYRFLPSPITVAEBOCCL 132
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 63 GATGEHVLALLADEVNPVSTVDADKQWTRQNLHVHVESGEGYRFVMPGATLDDDEFRRQ-L 121
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 133 DELGGAASAFAVVAAGSLPPGVADYQVRVADICRRSSTPLIIDTSGGGLQHTSS--GV 190
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 122 BEQVLEIESGAILVIVSGSLPPGVKVEKLTQLISAQKQICITIDSSDALTALALGDI 181
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 191 FLKASVRELRECEVSSLLTEPEQL-AAAHFLIDRGRAEVVVSIGSGALLATRHASHR 249
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 182 ELVVPNKELKLSALVNRD-LTQPDVDRKAQSLVSGKARRVVSIGPGALGISENCIQ 240
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 250 FSSIPMTAVSGVAGDANVAITVGLSRGWSLISVRLGNAAGAMLLTPTAACNRDV 309
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 241 VVPPVKSQSTVGAGDSWGMATLKLADQDASLEWVRFGVAAGSAATLNGTRLCSHDT 300
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 310 ERPF 313
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 301 QKIT 304
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

```

RESULT 10
US-10-369-493-851

```

; Sequence 851, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 851
; LENGTH: 309
; TYPE: PRT

```


PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 55963
LENGTH: 311
TYPE: PRT
ORGANISM: Enterobacter cloacae
US-10-282-122A-55963

Query Match 30.9%; Score 531.5; DB 4; Length 311;

Best Local Similarity 41.1%; Pred. No. 1,8e-41; Matches 125; Conservative 56; Mismatches 118; Indels 5; Gaps 4;

QY 14 IITLTMPALDITTSVDVVRPTKRCGAPRYDPGGGINVARIHVHVGCGSTALFPAGG 73
Db 5 IYTLSPSLDTSATITLPTQIYPEGKRCGAPRYDPGGGINVARIHVHVGCGSTALFPAGG 64
QY 74 STGSLMALLDGAGVPRFVPIPIASTRESFTVNESRTAKOYRFVLPGPSLTVAEQECLD 133
Db 65 ATGELHVSILADQVAVETVEAKDWTROMLHVHVGSSGEQYRFVWPGAKLSIDNEFRQ-LE 123
QY 134 ELRGAAAFVVAAGSLPPGVAADYQGVADICRRSTPLIDTSGGGLQH--ISSGVF 191
Db 124 EKVLTISGALLVSGSLPPGVSSTKLTALITMAQQRVIRCTVDSGSLALVPGMLE 183
QY 192 LUKASVRELRECVGSELTPEEQUL-AAAHELIDRGRAVNVVSLGSGCALATPASHRF 250
Db 184 LVKRNQKSLALVNRN-LTQPDVDTAAQELVTRCKARVVSLGPGQALAVDESSVQV 242
QY 251 SSIPMTAV-SCVAGADMVAITVGLSRGWSLKSIVRLGNAAGAMLLTPGTAA- 310
Db 243 VPPMKSGSTVAGDSMTGAMTLKLAQASILEMTRYGVAAGSAATINQTRLCISLADTQ 302
QY 311 RFFE 314
Db 303 KIIVD 306

RESULT 13

US-10-156-761-11226

Sequence 11226, Application US/10156761

Publication No. US20030119018A1

GENERAL INFORMATION:

APPLICANT: OMURA, SATOSHI

APPLICANT: IKEDA, HARUO

APPLICANT: ISHIKAWA, JUN

APPLICANT: HORIKAWA, HIROSHI

APPLICANT: SHIBA, TADAYOSHI

APPLICANT: SAKAKI, YOSHIYUKI

APPLICANT: HATTORI, MASAHIRA

TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES

FILE REFERENCE: 249-262

CURRENT APPLICATION NUMBER: US/10/156,761

PRIOR FILING DATE: 2002-05-29

PRIOR APPLICATION NUMBER: JP 2001-204089

PRIOR FILING DATE: 2001-05-30

PRIOR APPLICATION NUMBER: JP 2001-272697

PRIOR FILING DATE: 2001-08-02

NUMBER OF SEQ ID NOS: 15109

SEQ ID NO 11226

LENGTH: 317

TYPE: PRT

ORGANISM: Streptomyces avermitilis

US-10-156-761-11226

Query Match 20.6%; Score 355; DB 4; Length 317;
Best Local Similarity 32.8%; Pred. No. 9.3e-25;

Matches 106; Conservative 50; Mismatches 155; Indels 12; Gaps 8;

QY 14 IITLTMPALDITTSVDVVRPTKRCGAPRYDPGGGINVARIHVHVGCGSTALFPAGG 73
Db 2 ILYTTPNSLDRTREVSIDRGVIRATGEMDGGGVNVSRVAAGQTVAIPLCG 61
QY 74 STGSLMALLDGAGVPRFVPIPIASTRESFTVNESRTAKOYRFVLPGPSLTVAEQECLD 133
Db 62 APGAFVALLDQCIDVAPVAVAGATRSNISVAEP-DGTLTKINAPGELTPAAEELLE 120
QY 134 ELRG--AAASAFVVAAGSLPPGVAADYQGVADICRRSTPLIDTSGGGLQH--LGHISGV 190
Db 121 TVRQYRAGDITWTCCGSLPPGVAADYQGVADICRRSTPLIDTSGGGLQH--LGHISGV 180
QY 191 FLKASVRELRECVGSELTPEEQUL-AAAHELIDRGRAVNVVSLGSGCALATPASHRF 250
Db 181 DVKPNABELAEVGRPLATGADVAAHELRGIG-AGAVIASIGAGQGLVDSGTM-F 238
QY 251 SSIPMTAV-SCVAGADMVAITVGLSRGWSLKSIVRLGNAAGAMLLTPGTAA- 309
Db 239 GSARVDVNSVDVAGDSLSLGLFIAAGSGADALAS--AVAHGAAVQLPGSVMPKPADL 295
QY 310 RFFELAEPTVEGQDQYVHPI 332
Db 296 DP--SAVTVAEVPDRLTEPV 316

RESULT 14

US-10-156-761-11476

Sequence 11476, Application US/10156761

Publication No. US20030119018A1

GENERAL INFORMATION:

APPLICANT: OMURA, SATOSHI

APPLICANT: IKEDA, HARUO

APPLICANT: ISHIKAWA, JUN

APPLICANT: HORIKAWA, HIROSHI

APPLICANT: SHIBA, TADAYOSHI

APPLICANT: SAKAKI, YOSHIYUKI

APPLICANT: HATTORI, MASAHIRA

TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES

FILE REFERENCE: 249-262

CURRENT APPLICATION NUMBER: US/10/156,761

PRIOR FILING DATE: 2002-05-29

PRIOR APPLICATION NUMBER: JP 2001-204089

PRIOR FILING DATE: 2001-05-30

PRIOR APPLICATION NUMBER: JP 2001-272697

PRIOR FILING DATE: 2001-08-02

NUMBER OF SEQ ID NOS: 15109

SEQ ID NO 11476

LENGTH: 309

TYPE: PRT

ORGANISM: Streptomyces avermitilis

US-10-156-761-11476

Query Match 19.2%; Score 330; DB 4; Length 309;

Best Local Similarity 34.3%; Pred. No. 2.1e-22;

Matches 107; Conservative 44; Mismatches 141; Indels 20; Gaps 10;

QY 14 IITLTMPALDITTSVDVVRPTKRCGAPRYDPGGGINVARIHVHVGCGSTALFPAGG 73
Db 2 ILYTTPNSLDRTREVSIDRGVIRATGEMDGGGVNVSRVAAGQTVAIPLCG 61
QY 74 STGSLMALLDGAGVPRFVPIPIASTRESFTVNESRTAKOYRFVLPGPSLTVAEQECLD 133
Db 62 ATGELHVSILADQVAVETVEAKDWTROMLHVHVGSSGEQYRFVWPGAKLSIDNEFRQ-LE 123
QY 129 EOCIDELRGAAAFVVAAGSLPPGVAADYQGVADICRRSTPLIDTSGGGLQH-IS 187
Db 122 EAVGELIRASASA-----VALGSLPPGVAVAGVLAABAQAQVPLVDTSGEPLRGVA 177
QY 188 SCVFLKASVRELRECVGSELTPEEQUL-AAAHELIDRGRAVNVVSLGSGCALATPASHRF 247
Db 178 ARPDIVPNABELAEELTGSN---EPWQ--ATRDARRRG-ANTVVASIGAGG-LIAVVAEG 230

```

Oy      248 HRFSSIPMTAVSG--VGA GDANVAATVGLSRGWSLKSIVRIGNAGAAAMLLTPGAACN 305
Db      231 H-WRATPEHRRVKGNTGTGAGDSVVAAGLSGLVQDLPMPKRLAAVAALSAATVLSFVAGEFD 289
Oy      306 RDDVERFFELAA 317
Db      290 RRTVEELDRVA 301

RESULT 15
US-10-282-122A-45630
: Sequence 45630, Application US/10282122A
: Publication No. US20040029129A1
: GENERAL INFORMATION:
: APPLICANT: Wang, Liangu
: APPLICANT: Zamudio, Carlos
: APPLICANT: Malone, Cheryl
: APPLICANT: Haselbeck, Robert
: APPLICANT: Ohlsen, Kari
: APPLICANT: Zykiedg, Judith
: APPLICANT: Wali, Daniel
: APPLICANT: Trawick, John
: APPLICANT: Carr, Grant
: APPLICANT: Yamamoto, Robert
: APPLICANT: Foreyth, R.
: APPLICANT: Xu, H.
: TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
: FILE REFERENCE: ELITRA.034A
: CURRENT APPLICATION NUMBER: US/10/282,122A
: CURRENT FILING DATE: 2003-02-20
: PRIOR APPLICATION NUMBER: 60/191,078
: PRIOR FILING DATE: 2000-03-21
: PRIOR APPLICATION NUMBER: 60/206,848
: PRIOR FILING DATE: 2000-05-23
: PRIOR APPLICATION NUMBER: 60/207,727
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: 60/230,335
: PRIOR FILING DATE: 2000-09-06
: PRIOR APPLICATION NUMBER: 60/230,347
: PRIOR FILING DATE: 2000-09-09
: PRIOR APPLICATION NUMBER: 60/242,578
: PRIOR FILING DATE: 2000-10-23
: PRIOR APPLICATION NUMBER: 60/253,625
: PRIOR FILING DATE: 2000-11-27
: PRIOR APPLICATION NUMBER: 60/257,931
: PRIOR FILING DATE: 2000-12-22
: PRIOR APPLICATION NUMBER: 60/267,636
: PRIOR FILING DATE: 2001-02-09
: PRIOR APPLICATION NUMBER: 60/269,308
: PRIOR FILING DATE: 2001-02-16
: Remaining Prior Application data removed - See File Wrapper or PAM.
: NUMBER OF SEQ ID NOS: 78614
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 45630
: LENGTH: 303
: TYPE: PRT
: ORGANISM: Bacillus anthracis
US-10-282-122A-45630

Query Match      18.1%; Score 311; DB 4; Length 303;
Best Local Similarity 29.0%; Pzed. No. 1.3e-20;
Matches      89; Conservative 60; Mismatches 132; Indels 26; Gaps 8

Oy      14 IITLTMMNALDITTSVD-----VVRPEKMGCGAPRYDGGGGGVNARIYVHGGCSTRA 67
Db      2 IYTYTLNPSIDVVQVNSLIDGAVNRAEKOMF-----PGKGINVSVLHRLGKENVYA 55

Oy      68 LFPYAGSGTSLMALIDGAGVFRYIPIASTRESFTV---NESRTAKORYFVLPGPSLT 124
Db      56 LGFTGCGFGRFKIVDLQGTREGVITNRYVDGDSRINVKIKGGEFTELNQG-----GPSVT 109

Oy      125 VAEQEQCLDELKGAASAAFFVVASGLPGCVADYIQRVADICRRSSTPLIIDTSGGGGLQ 184

```

[illegible]

Search completed: March 23, 2006, 06:28:10
Job time : 72.2896 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceeleration Ltd.

OM protein - protein search, using sw model

Run on: March 23, 2006, 06:17:23 ; Search time 8.13514 Seconds
(without alignments)
1192.746 Million cell updates/sec

Title: US-10-617-038-22

Perfect score: 1120
Sequence: 1 MTEPAWDEGKPRITITLMN.....TEVGQDQYVHPIVNPASP 339

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 169630 seqs, 2862289 residues

Total number of hits satisfying chosen parameters: 169630

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA_New:*
1: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	283	16.5	306	US-10-485-517-185	Sequence 185, App
2	168	9.8	488	US-11-087-099-303	Sequence 303, App
3	165.5	9.6	323	US-11-132-864-7	Sequence 7, Appl
4	164.5	9.6	470	US-11-087-099-11203	Sequence 11203, A
5	164.5	9.6	496	US-11-087-099-6042	Sequence 6042, Ap
6	161	9.4	345	US-11-096-686-11061	Sequence 11061, A
7	145	8.4	473	US-11-087-099-12426	Sequence 12426, A
8	142.5	8.3	480	US-11-087-099-6193	Sequence 6193, Ap
9	140.5	8.2	361	US-11-087-099-7506	Sequence 7506, Ap
10	131	7.6	369	US-11-096-568A-6281	Sequence 6281, Ap
11	131	7.6	369	US-11-096-568A-6280	Sequence 6280, Ap
12	129.5	7.5	480	US-11-087-099-6825	Sequence 6825, Ap
13	128.5	7.5	176	US-11-087-099-7668	Sequence 7668, Ap
14	122	7.0	320	US-10-467-657-3254	Sequence 3254, Ap
15	121	7.0	302	US-10-793-626-3062	Sequence 3062, Ap
16	120	7.0	187	US-11-087-099-11101	Sequence 11101, A
17	116	6.7	212	US-11-096-568A-31212	Sequence 6282, Ap
18	115.5	6.7	378	US-11-096-568A-31212	Sequence 31212, A
19	115.5	6.7	379	US-11-096-568A-31211	Sequence 31211, A
20	115.5	6.7	705	US-11-096-568A-31090	Sequence 31090, A
21	115.5	6.7	741	US-11-096-568A-31089	Sequence 31089, A
22	115.5	6.7	750	US-11-096-568A-31088	Sequence 31088, A
23	114	6.6	397	US-11-132-864-28	Sequence 28, Appl
24	113.5	6.6	428	US-11-096-568A-12725	Sequence 12725, A
25	112	6.5	261	US-11-096-568A-31213	Sequence 31213, A

26	111.5	6.5	316	US-11-087-099-9229	Sequence 9229, Ap
27	110.5	6.4	477	US-10-520-820-13	Sequence 11, Appl
28	109.5	6.4	323	US-11-087-099-1760	Sequence 1760, Ap
29	108.5	6.3	307	US-11-087-099-4308	Sequence 4308, Ap
30	108.5	6.3	310	US-11-087-099-10267	Sequence 10267, A
31	108.5	6.3	322	US-11-096-568A-10266	Sequence 10266, A
32	108.5	6.3	359	US-11-096-568A-10265	Sequence 10265, A
33	106.5	6.2	117	US-09-995-493-40	Sequence 40, Appl
34	106.5	6.2	399	US-11-087-099-2847	Sequence 2847, Ap
35	105.5	6.1	313	US-11-098-686-11025	Sequence 11025, A
36	105	6.1	319	US-11-087-099-10858	Sequence 10858, A
37	105	6.1	378	US-11-096-568A-12726	Sequence 12726, A
38	101	5.9	319	US-10-793-626-786	Sequence 786, App
39	98.5	5.7	388	US-11-132-864-41	Sequence 41, Appl
40	98.5	5.7	5291	US-11-052-554A-281	Sequence 281, App
41	98	5.7	353	US-11-132-864-29	Sequence 29, Appl
42	98	5.7	380	US-11-132-864-30	Sequence 30, Appl
43	98	5.7	1170	US-11-046-456-28	Sequence 28, Appl
44	98	5.7	1170	US-11-046-644-28	Sequence 28, Appl
45	98	5.7	6893	US-11-205-109-14	Sequence 14, Appl

ALIGNMENTS

```
RESULT 1
US-10-485-517-185
; Sequence 185, Application US/10485517
; Publication No. US20050256299A1
; GENERAL INFORMATION:
; APPLICANT: University of Sheffield
; APPLICANT: Biosynexus Incorporated
; APPLICANT: Foster, Simon
; APPLICANT: Mond, James
; TITLE OR INVENTION: Antigenic Polypeptides
; FILE REFERENCE: P10629W0
; CURRENT APPLICATION NUMBER: US/10/485,517
; CURRENT FILING DATE: 2004-02-02
; PRIOR APPLICATION NUMBER: GB 0118625.9
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: GB 0200349.9
; PRIOR FILING DATE: 2002-01-09
; NUMBER OF SEQ ID NOS: 424
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 185
; LENGTH: 306
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
; US-10-485-517-185

Query Match      16.5%; Score 283; DB 6; Length 306;
Best Local Similarity 26.6%; Pred. No. 2,5e-16;
Matches 81; Conservative 66; Mismatches 139; Indels 18; Gaps 7;

14 IITLVNPAIDITTSVDVVRPTKRCGAPRYPGGGGINVARIHVLAGCSTALFPAGG 73
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||
2 IYVTFRPSIDVYFTDFKIDGLNRAFATYKFKAGGKINVSRLKTLVDSTALPAGG 61
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||
74 STGSLMALIGDAGVPRVPIAASSTREFTVN-ESSTAQYRFVLPGBSLVAEDQCL 132
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||
62 FPGFFIIDTLNNSIQSNFIEVDEDTF---INVKTKTGETEINAGPHITSTGFQSL 117
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||
133 DELRGAAASAFVVASGSLPFGVAADYVORVADICRSSTPLIDTSGGGLQHTSSV-- 190
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||
118 QQIQN-TTSEDYIVVAGSVSSIPSDAYIAQITTAQTGAKLVDAE---KEIAESVLP 172
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||
191 ---FLKASVRELRCEVSELTPEQGLAAHEIDRGRAEVVVVSGGQALATPHAS 247
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||
173 YHPLFKPNDELVEWMENTVNSDTVIKYGRLLVDKG-AQSVIVSGGGAIIYDKSIS 231
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||
248 HRSSSIMTVTSVGVGADANVAATITVGLSRGWSIKSVRLGNAAGAMLLTPGAAACNRD 307
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||
232 IKAVNPOGKVVNTVGSDDSTVAGVAGIAGLTIEKAFOQAVACGTAFDEDLA--TRD 289
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||
```

Qy 308 DVER 311
Db 290 AIEK 293

RESULT 2

US-11-087-099-303
; Sequence 303, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 303
; LENGTH: 488
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-11-087-099-303

Query Match 9.8%; Score 168; DB 7; Length 488;
Best Local Similarity 22.9%; Pred. No. 1.9e-06;
Matches 85; Conservative 56; Mismatches 156; Indels 74; Gaps 14;

Qy 11 KPRITITLMPALDITTSVDVVRP-----TEKRCGAP--RYDPGGGGINVARIW 59
Db 83 KPIVSTLGNLCVDIVLSVHSLPPPSRGERKALMDELMSPPDKKYWEAGNCMAIAAA 142
Qy 60 VLGGCSTALPPAGST--GSLIMALLGDAGVPRVPIPIASTRESFTYNESRTAKQYFVL 118
Db 143 RLGIHCAVAGHVGDEIYGEFLDVLHBEGL--GTVALNGGTNEKDTSSFCETL--ICWVL 198
Qy 119 RGP-----SLTAEBOEQL-----DELGAASAFAFVVASGLSPGVADYQYRVA 164
Db 199 VDPQRHGFSGRADPKKEBPASWITTDSDVKMIRSKVLCFCGYDPDDPSFIMSTI 258
Qy 165 DICRSSTPLILDTSGGLQHISSGV-----FLTKASVRLRCVSGSELITPE-- 213
Db 259 DYAKVGTATFFDGPGRG--KSLSGTPTDERALLHFLRMSVYL-----LITSEVE 308
Qy 214 -----QLAAHLEIDRGR-AEYVVVSLSGGALLATRHASHRSSIIPMTAVSGVAGD 265
Db 309 ALTGIRNPVAGOEILNKGKTKWIVMGPKGSILVTKSSVAPAPAKYEVVDVTYCGD 368
Qy 266 AMVAATVVGSRGWSLTKSVTLGNAAGAMLLTPGTA-----ACNRDVER 311
Db 369 SFVAALAGYIRNPLVNTLTITANAGATAMGCGAGRNVAKRHQVVDLMKASKLNDDEK 428
Qy 312 FFE-LAAEPT 321
Db 429 FFEQLLENSE 439

RESULT 3

US-11-132-864-7
; Sequence 7, Application US/11132864
; Publication No. US20050289670A1
; GENERAL INFORMATION:
; APPLICANT: Jimui Shi
; APPLICANT: David Erci
; APPLICANT: Lisa Hagen
; APPLICANT: Hongyu Wang
; TITLE OF INVENTION: Plant Myo-Inositol Kinase
; FILE REFERENCE: 035718/291638
; CURRENT APPLICATION NUMBER: US/11/132,864
; CURRENT FILING DATE: 2005-05-19
; PRIOR APPLICATION NUMBER: 60/573,000
; PRIOR FILING DATE: 2004-05-20
; NUMBER OF SEQ ID NOS: 48

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 323
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Pfam consensus for pfkb family
US-11-132-864-7

Query Match 9.6%; Score 165.5; DB 7; Length 323;
Best Local Similarity 28.1%; Pred. No. 1.8e-06;
Matches 82; Conservative 33; Mismatches 128; Indels 49; Gaps 12;

Qy 47 PGGGGINVARIIVHLGGCSTALPPAGST-----GSLMALLDAGVPRVPIPIASTRE 101
Db 35 PGAGANVAVVALRLGIPSGYKVAFLGKVGDDDFEFLLELLKKEGVDYVVDYVDEGRT 94
Qy 102 SFTV-----NESRTKQYFVLPGSLTVAE--QEGLDLELRGAASAFAFVVASGLSPG 154
Db 95 GLALVLDVGDGGRITVYRGA--NADLTLELDELDLEADITLHSGISVL--LPEP 149
Qy 155 VAADYQYRVADICRRS-----TPILDTSG--GGLQHISGVFLKASVRELRECV 204
Db 150 LPETLEALAEAKAGKISFDPNLRDPLMSDEALEVLELLPLADILKPNBEL-- 205
Qy 205 GSELIT-----EPQLAAHLEIDRGR-AEYVVVSLSGGALLATRHASHRSSIP-MTA 257
Db 206 --SILTGLKGEVEEALAKLAKAVTKLVVVTGADGALLVTGGEVNPVAVPKVY 263
Qy 258 VSGVGADAMVAALITVGLSR-----GWSLTKSVRLGNAAGAMLLTPG 300
Db 264 VDTTGADAVAGLAGLATDSDNTQLDGDLEALRFANMAALVVOCKG 315

RESULT 4

US-11-087-099-11203
; Sequence 11203, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 11203
; LENGTH: 470
; TYPE: PRT
; ORGANISM: Oryza sativa (japonica cultivar-group)
US-11-087-099-11203

Query Match 9.6%; Score 164.5; DB 7; Length 470;
Best Local Similarity 22.6%; Pred. No. 3.6e-06;
Matches 84; Conservative 53; Mismatches 181; Indels 53; Gaps 12;

Qy 8 DEGPRIITLTMALDITTSVDVVRP-----EKRCGAP--RYDPGGGGINVAR 56
Db 65 EBKERTDLATLGNLCVDIVLSVPLPPAREEREAUMLRAASPPQKFWBAGNCMLAF 124
Qy 57 IVAVLG--GCGSTALPPAGSTGSLMALLDAGVPRVPIPIASTRESFTYNESRTAKQY 115
Db 125 AAARLGRCSLGHVBEIYKFLDVLBEELG--SVGMLNDSGACRMAVETLLCW 182
Qy 116 FVLPGSLITVAEOECID-----ELRGAASAFAFVVASGLSPGVADYQYRV 163
Db 183 LVDPFGHGFSGRADSDERAFSWIKRLPETGTALHNSKILCNCVAFDELPDYISSA 242
Qy 164 ADICRSSTPLILDTGGG-----LQHT--SGGVFLKASVRELRECVSEL 208
Db 243 IDCALDAGTAVFPDPGRKSLHGTLDDEORALEHSLRSDVLLTSDENAE-----SL 295
Qy 209 LTEREQLAAHLEIDRG--BAEYVVVSLSGGALLATRHASHRSSIIPMTAVSGVAGDAM 267

Db 236 TNNPPIAGGELKRGIRTKWVYIKMSKSGSIMVTSKAVSABSEFKIDVDVITGCGDSF 355
Qy 268 VAAITVGLSRGMSLTKSVRLGNAAGAMLLTPTGAACNRDVERFEELAAEPTVEGQDQY 327
Db 356 TAAIAFGFLHNLPAVSTLTLLANAVGAATATGCG-AGRVVAHLDKVLQLLRE-SNINEDDT 413
Qy 328 VMHPIVNPEAS 338
Db 414 PMSSELI--EAS 422

RESULT 5

US-11-087-099-6042
; Sequence 6042, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087, 099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 6042
; LENGTH: 496
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(496)
; OTHER INFORMATION: unsure at all xaa locations
US-11-087-099-6042

Query Match 9.6%; Score 164.5; DB 7; Length 496;
Best Local Similarity 22.6%; Pred. No. 3.9e-06;
Matches 84; Conservative 53; Mismatches 181; Indels 53; Gaps 12;

Qy 8 DEGRPRITTLTMNPALDITTSVDVVRPT-----EKMRGAP--RYDPGGGGINVAR 56
Db 87 EEGKETDLATLGNLCVDVLSVLPQLPPAPRERERAYMERLAASPPDQKFWBAGNCMLAF 146
Qy 57 IVHVLG-GCSTALPPAGGTSLLMALLGDAGVPPRVIPPIAASRESTVNESTRAKOYR 115
Db 147 AARLGRGCTLGHVGEIYKFLDLVEEGI--SVGMLDNSDSCACRAVAYETLLCWV 204
Qy 116 FVLPGSLTVAEOECQD-----ELRGAASAFAVVASGLPPGVADYQYQV 163
Db 205 LVDPFGHGFSCRADFSDEPAFWSIRKLPATETKAIHHSKILFCNGIAPDELPDVISSA 264
Qy 164 ADICRSSTPLILDTSGGG-----LQHI--SSGVFLKASVRELRECVGSEL 208
Db 265 IDCAIDAGTAVFPDPGRGKSLHGTDEGRALHSLRLSDVLLTSDAE-----SL 317
Qy 209 LTEPEQLAAAHLEIDRG-RAEVVVSIGSGALLATRHASHRFSSIPMTAVSGVAGADM 267
Db 318 TNNPPIAGGELKRGIRTKWVYIKMSKSGSIMVTSKAVSABSEFKIDVDVITGCGDSF 377
Qy 268 VAAITVGLSRGMSLTKSVRLGNAAGAMLLTPTGAACNRDVERFEELAAEPTVEGQDQY 327
Db 378 TAAIAFGFLHNLPAVSTLTLLANAVGAATATGCG-AGRVVAHLDKVLQLLRE-SNINEDDT 435
Qy 328 VMHPIVNPEAS 338
Db 436 PMSSELI--EAS 444

RESULT 6

US-11-098-686-11061
; Sequence 11061, Application US/11098686
; Publication No. US20060024696A1
; GENERAL INFORMATION:
; APPLICANT: Kapur, Vivek and Gebhart, Connie J.
; TITLE OF INVENTION: NUCLEIC ACID AND POLYPEPTIDE SEQUENCES
; TITLE OF INVENTION: FROM LAWSONIA INTRACELLULARIS AND METHODS OF USING

; FILE REFERENCE: 09531-128001
; CURRENT APPLICATION NUMBER: US/11/098, 686
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: PCT/US03/31318
; PRIOR FILING DATE: 2003-10-01
; PRIOR APPLICATION NUMBER: US 60/416,395
; NUMBER OF SEQ ID NOS: 11433
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11061
; LENGTH: 345
; TYPE: PRT
; ORGANISM: Lawsonia intracellularis
US-11-098-686-11061

Query Match 9.4%; Score 161; DB 7; Length 345;
Best Local Similarity 21.4%; Pred. No. 4.7e-06;
Matches 72; Conservative 59; Mismatches 160; Indels 46; Gaps 12;

Qy 11 KPRIT--LTMPALDITTSVDVVRPTERKRCGAPRYDPGGGGINVARIHVVLGCGSTA 67
Db 21 KPVLVGDIMLDNVLIGVSDRISPEAPVPIKTIENKQSLGAG-IVARSIALLGKVTI 79
Qy 68 LFPAG-GSTGSLMALIGDAGVPPRVIPIA-----ASTR-----ESFTVNE 107
Db 80 IGAVGQDQSGEKIDLLSTRGILSSITTFANRQTVTRVMAHQMIRLDHEESTSYN- 138
Qy 108 SRTAKOYRFLPGSLTVAEOE-QCDELGAASAFAVVASGLPPGVADYQYQVADI 166
Db 139 --SKELAMVLSNEKYSOHEIILISDYNKGLVSKFELG-----FQDIL-L 182
Qy 167 CRSSSTPLILDTSGGGLQH--ISSGVFLKASVRELRECVGS-ELTEPEQLAAHLEID 223
Db 183 AKNNMAKVLIDPRPCNIMYVALCKNIFALIPNTEKEGCGAMATSGQPELLAAGHTIMK 242
Qy 224 RGRAEVVVSIGSG-ALLATRHASHRFSSIPMTAVSGVAGADMVAITVGLSRGMSLI 282
Db 243 LLSFKHLITLIGDSGMLFLSPKKIWHIPVGRDVPDVTGAGDTVATFGIALSAGIDPL 302
Qy 283 KSVRLGNAAGAMLLTPTGAACNRDVERFEELAAEP 319
Db 303 ISAILANYAAGVAVSQVGTAVSPDELEAITSLPOP 339

RESULT 7

US-11-087-099-12426
; Sequence 12426, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087, 099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 12426
; LENGTH: 473
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(473)
; OTHER INFORMATION: unsure at all xaa locations
US-11-087-099-12426

Query Match 8.4%; Score 145; DB 7; Length 473;
Best Local Similarity 21.3%; Pred. No. 0.00016;
Matches 81; Conservative 60; Mismatches 169; Indels 70; Gaps 15;

Qy 4 PAAMDEKPRITTLTMNPALDITTSVDVVRPTERKRCGA-----PRIDPGGGGI 52
Db 64 PMSW---KHVDVATLSNLCDIVLVNLPQLPSPPLQRKAFMDRLAASPPDKKYWAGGNC 120


```

Qy 91 RVIPAASTRESFTYNESRTAKQYRFV-LPSPSLTV---AEQECLELRCGAASAATV 146
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 134 DNLAIVASAAATGHAIVMLQNSQNSIYVIGGANLSCWPSLPRQHLD---LVAQAGIVL 189
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 147 ASGSLPQVADYQVADICRRSSTPLIDTSS--GGL-QHISGVFLKASVRELREC 203
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 190 LQREIPDAVNA---QVAQAANKAGVPVPLDAGGMDGLPQLINFVDILSPNETELARL 245
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 204 VG--SELLTEPEOLA-AAHELIDRGRAVVVSLGSGALLAT--RHASHRFPSSIPMTAV 258
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 246 TGMPTSEFEELIQOALKCHEM---GAKQVLVKLGKHSALFVBESEKTIQOPAILAKTV 301
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 259 SGVGAGDAMVAATVGLSRGWSLKSIVRLGNAGAAAMLLTPGTA 302
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 302 DTTGAGDTFTAAFAVALVEGSKKECLRFPAAAACLCVQVKGAS 345
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 11
US-11-096-568A-6280
; Sequence 6280, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 6280
; LENGTH: 369
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(369)
; OTHER INFORMATION: Ceres Seq. ID no. 14314594
US-11-096-568A-6280

```

```

Query Match 7.6%; Score 131; DB 7; Length 369;
Best Local Similarity 26.1%; Pred. No. 0.0017;
Matches 74; Conservative 37; Mismatches 133; Indels 40; Gaps 13;

Qy 47 PGCGGINVARIIVHLVG-----CSTPL-FP-----AGSTGSLMLLGDAGVPF 90
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 82 PGBETTLAARSGQTAGCKGANKATCSAKLAVPTVQVGDADAGRLVATGKGGVRL 141
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 91 RVIPAASTRESFTYNESRTAKQYRFV-LPSPSLTV---AEQECLELRCGAASAATV 146
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 142 DNLAIVASAAATGHAIVMLQNSQNSIYVIGGANLSCWPSLPRQHLD---LVAQAGIVL 197
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 147 ASGSLPQVADYQVADICRRSSTPLIDTSS--GGL-QHISGVFLKASVRELREC 203
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 198 LQREIPDAVNA---QVAQAANKAGVPVPLDAGGMDGLPQLINFVDILSPNETELARL 253
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 204 VG--SELLTEPEOLA-AAHELIDRGRAVVVSLGSGALLAT--RHASHRFPSSIPMTAV 258
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 254 TGMPTSEFEELIQOALKCHEM---GAKQVLVKLGKHSALFVBESEKTIQOPAILAKTV 309
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 259 SGVGAGDAMVAATVGLSRGWSLKSIVRLGNAGAAAMLLTPGTA 302
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 310 DTTGAGDTFTAAFAVALVEGSKKECLRFPAAAACLCVQVKGAS 353
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 12
US-11-087-099-6825
; Sequence 6825, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP

```

```

; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 6825
; LENGTH: 480
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(480)
; OTHER INFORMATION: unsure at all Xaa locations
US-11-087-099-6825

```

```

Query Match 7.5%; Score 129.5; DB 7; Length 480;
Best Local Similarity 26.7%; Pred. No. 0.0031;
Matches 48; Conservative 28; Mismatches 77; Indels 27; Gaps 7;

Qy 166 ICRSSSTPLIDTSGCG-----LQHT--SSGVFLKASVRELRECSELLT 210
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 257 IC--SGTAVFPDPGRKSLHGTGLGROALEHLRLSDVLLTDEAE-----SLTN 307
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 211 EPEQLAAHELIDRG-RAEVVVSLSGSGALLATRHASHRFPSSIPMTAVSGVAGDAMVA 269
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 308 IRNVQAGQELKRGITKQVIVKMGSRGSMITKNTISCAPAKIDVDTVGGGDSFTA 367
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 270 AITVGLSRGWSLKSIVRLGNAGAAAMLLTPGTAACNRDVERFPFELAAEPTVGQDYVW 329
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 368 AIAFGFLHDLPAVNTLTLANAVGAATATGCG-AGRNVARLDKVLQVRE-ADLNEDITLW 425
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 13
US-11-087-099-7668
; Sequence 7668, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 7668
; LENGTH: 176
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(176)
; OTHER INFORMATION: unsure at all Xaa locations
US-11-087-099-7668

```

```

Query Match 7.5%; Score 128.5; DB 7; Length 176;
Best Local Similarity 33.3%; Pred. No. 0.001;
Matches 36; Conservative 18; Mismatches 51; Indels 3; Gaps 3;

Qy 215 LAAHELIDRG-RAEVVVSLSGSGALLATRHASHRFPSSIPMTAVSGVAGDAMVAATV 273
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 47 IQAGQELKRGITKQVIVKMGSRGSMITKNTISCAPAKIDVDTVGGGDSFTAIAF 106
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 274 GLSRGWSLKSIVRLGNAGAAAMLLTPGTAACNRDVERFPFELAAEPT 320
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 107 GFLHNLPAVSTLTLANAVGAATATGCG-AVLSKTAVNSFSDRLVHVPT 153
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 14
US-10-467-657-3254
; Sequence 3254, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega

```


GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: March 23, 2006, 04:56:29 ; Search time 181.055 Seconds
(without alignments)
1652.628 Million cell updates/sec

Title: US-10-617-038-23

Sequence: 1 VLMTAAADYRRSPRVFRD.....LEVTSRWIAGENPEYPTGL 681

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: A_Geneseq_21:*
2: Geneseqp1980s:*
3: Geneseqp1990s:*
4: Geneseqp2000s:*
5: Geneseqp2001s:*
6: Geneseqp2002s:*
7: Geneseqp2003s:*
8: Geneseqp2004s:*
9: Geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3508	100.0	681	8 ADI37302	Adi37302 M. tuberc
2	547.5	15.6	226	9 ABM91252	Abm91252 M. xanthu
3	514	14.7	217	9 ADU49950	Adu49950 Methanosa
4	464	13.2	220	9 AEB36491	Aeb36491 L. pneumo
5	463	13.2	216	9 AEB39892	Aeb39892 L. pneumo
6	431	12.3	443	9 AEB95281	Aeb95281 M. tubercu
7	431	12.3	443	9 AEB93950	Aeb93950 Mycobacte
8	420.5	12.0	233	9 ADU49942	Adu49942 Meloidogy
9	420.5	12.0	233	9 AEB95272	Aeb95272 M. javanic
10	420.5	12.0	233	9 AEB93956	Aeb93956 Root-knot
11	417.5	11.9	229	9 ABM96175	Abm96175 M. xanthu
12	417.5	11.9	233	9 ADU49941	Adu49941 Meloidogy
13	417.5	11.9	233	9 AEB95271	Aeb95271 M. incogni
14	417.5	11.9	233	9 AEB93944	Aeb93944 Southern
15	415.5	11.8	289	7 ADB74647	Adb74647 Mycobacte
16	414.5	11.8	257	9 ADU49952	Adu49952 Azotobact
17	413.5	11.8	235	9 ADU49949	Adu49949 Mesorhizo
18	412.5	11.8	226	9 ADU49951	Adu49951 Novosphin
19	406	11.6	236	9 ADU49944	Adu49944 Sinorhizo
20	401.5	11.4	229	9 ADU49943	Adu49943 Heteroder
21	401.5	11.4	229	9 AEB95273	Aeb95273 H. glycine
22	401.5	11.4	229	9 AEB93945	Aeb93945 Root-knot
23	401.5	11.4	229	9 AEB93946	Aeb93946 Soybean c
24	346.5	9.9	227	9 ABM93647	Abm93647 M. xanthu

25	221	6.3	419	1 AAP82072	Aap82072 Esterase
26	157	4.5	226	2 AAW98370	Aaw98370 H. pylori
27	156.5	4.5	4635	8 ADX56098	Adx56098 Streptomy
28	152	4.3	234	2 AAY11078	Aay11078 H. pylori
29	149	4.2	2552	6 ABU33748	Abu33748 Protein e
30	147.5	4.2	1650	7 ABO73052	AbO73052 Pseudomon
31	147	4.2	996	6 ABU22241	Abu22241 Protein e
32	145	4.1	3192	4 AAE10128	Aae10128 Streptomy
33	144	4.1	1741	6 ABM22363	Abm22363 Protein e
34	143.5	4.1	2214	9 ABM96789	Abm96789 M. xanthu
35	143	4.1	836	6 ABU41360	Abu41360 Protein e
36	143	4.1	8360	6 AAE35499	Aae35499 Streptomy
37	141.5	4.0	4799	8 ADX56094	Adx56094 Streptomy
38	141	4.0	4105	5 ABE62762	Abp62762 S. roseos
39	141	4.0	7338	7 ADJ72174	Adj72174 Streptomy
40	140	4.0	168	8 ADN26380	Adn26380 Bacteri
41	139.5	4.0	2130	8 ADQ74677	Adq74677 Streptomy
42	138	3.9	1620	8 ADQ74676	Adq74676 Streptomy
43	136.5	3.9	1340	6 ABU19856	Abu19856 Protein e
44	136	3.9	685	7 ABO77977	AbO77977 Pseudomon
45	136	3.9	846	8 ADN26449	Adn26449 Bacteri

ALIGNMENTS

RESULT 1
ADI37302 standard; protein; 681 AA.
ADI37302;
22-APR-2004 (first entry)
M. tuberculosis low oxygen induced antigen Rv2030c SEQ ID NO:23.
mycobacterial infection; vaccine; tuberculosis;
mycobacterium tuberculosis; immunisation; antibacterial; gene therapy;
low oxygen induced antigen.
Mycobacterium tuberculosis.
WO2004006952-A2.
22-JAN-2004.
08-JUL-2003; 2003WO-DK00477.
13-JUL-2002; 2002DK-00001098.
(STAT-) STATENS SERUM INST.
Andersen P, Rosenkrands I, Stryhn A;
WPI; 2004-122778/12.
N-PSDB; ADI37347.
Use of one or more polypeptides or their fragments, which are expressed during the latent stage of the mycobacterial infection, and/or nucleic acids encoding the polypeptides, for a therapeutic vaccine against tuberculosis.
Claim 3; SEQ ID NO 23; 76pp; English.
The present invention describes polypeptides or their fragments, which are expressed during the latent stage of a mycobacterial infection, and/or nucleic acids encoding the polypeptides, which are useful for creating a therapeutic vaccine against tuberculosis. Also described: (1) a therapeutic vaccine against tuberculosis comprising one or more polypeptides; (2) a method for treating an animal, including a human being, with tuberculosis caused by virulent mycobacteria, e.g. by Mycobacterium tuberculosis, M. africanum or M. bovis; (3) a method for immunising an animal, including a human being, against tuberculosis

CC caused by virulent mycobacteria; (4) a method of diagnosing tuberculosis
CC caused by virulent mycobacteria in an animal, including a human being;
CC (5) a method for diagnosing previous or ongoing infection with a virulent
CC mycobacterium; and (6) a method of diagnosing Mycobacterium tuberculosis
CC infection in a subject. The polypeptides have antibacterial activities,
CC and can be used in vaccines and in gene therapy. The polypeptides are
CC useful for the manufacture of a therapeutic vaccine for treating an
CC individual who is infected by a virulent mycobacterium, e.g. M.
CC tuberculosis, and who is not vaccinated with BCG against tuberculosis.
CC The present sequence represents a low oxygen induced antigen, which is
CC used in the exemplification of the present invention.

XX
SQ Sequence 681 AA;

Query Match 100.0%; Score 3508; DB 8; Length 681;

Best Local Similarity 100.0%; Pred. No. 0;
Matches 681; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLMTAAADVTRSRPRVRPRDRREGAVLAELAAVROPDVIYGLARGGLPVAMEVAAA 60
DB 1 VLMTAAADVTRSRPRVRPRDRREGAVLAELAAVROPDVIYGLARGGLPVAMEVAAA 60
QY 61 LHAFLDAFVVRKLGAPGHDEPAVGALASGGRVNVNDVVRGLRTTPOOLRDIAREGREL 120
DB 61 LHAFLDAFVVRKLGAPGHDEPAVGALASGGRVNVNDVVRGLRTTPOOLRDIAREGREL 120
QY 121 LRRESAYRGERPPTDITGKTVIVDDGLATGASMFPAVQALRDAQPAQIVAVPAPEST 180
DB 121 LRRESAYRGERPPTDITGKTVIVDDGLATGASMFPAVQALRDAQPAQIVAVPAPEST 180
QY 181 CREPAGLVDDVVCATMTPTPLAVGESFWDPROQVDEVRLLAFTPTGSPSRRAASTAA 240
DB 181 CREPAGLVDDVVCATMTPTPLAVGESFWDPROQVDEVRLLAFTPTGSPSRRAASTAA 240
QY 241 DVLRRVAIDAPGVPTEHEVLAELVGDARIVLIGESSHGTHEFYQARAAMTOMLIEEGFG 300
DB 241 DVLRRVAIDAPGVPTEHEVLAELVGDARIVLIGESSHGTHEFYQARAAMTOMLIEEGFG 300
QY 301 AVAAEADWMPDAYRVNRYVRGLGEDTNADEALSGFERPPAMWMTVVRDFVEMLRTRNOR 360
DB 301 AVAAEADWMPDAYRVNRYVRGLGEDTNADEALSGFERPPAMWMTVVRDFVEMLRTRNOR 360
QY 361 YESGALFQAGFYGLDLYSHRSIOEVI SYLDKVPRAAARARAYACFDHACADDDGQAYG 420
DB 361 YESGALFQAGFYGLDLYSHRSIOEVI SYLDKVPRAAARARAYACFDHACADDDGQAYG 420
QY 421 FAAAFAFGSPGSCERAVBOLVDVQRNALAYARODGLAEDELFFVAQONAOCTRADAEVYVRA 480
DB 421 FAAAFAFGSPGSCERAVBOLVDVQRNALAYARODGLAEDELFFVAQONAOCTRADAEVYVRA 480
QY 481 MFSGRVTSWMLRDQMAQTGLSLTLHLDRHLDAPPAIYVAHNSHVGDARATEVMADGQ 540
DB 481 MFSGRVTSWMLRDQMAQTGLSLTLHLDRHLDAPPAIYVAHNSHVGDARATEVMADGQ 540
QY 541 LITGQIVREKYGDSRSIGSTTYGTTAASEWGGIQRKAVRALGSEVEELFHQTADS 600
DB 541 LITGQIVREKYGDSRSIGSTTYGTTAASEWGGIQRKAVRALGSEVEELFHQTADS 600
QY 601 FLVSARLSRDAABPLDVRLGRATGVVYLPTEROSHLYLHVRPADQPDAMITHIDQTALS 660
DB 601 FLVSARLSRDAABPLDVRLGRATGVVYLPTEROSHLYLHVRPADQPDAMITHIDQTALS 660
QY 661 PLEVTSRWINGENPETYPTGI 681
DB 661 PLEVTSRWINGENPETYPTGI 681

RESULT 2
ID ABM91252 standard; protein; 226 AA.
XX
AC ABM91252;
XX

DT 02-JUN-2005 (first entry)

XX M. xanthus protein sequence, seq id 10451.

DE Transgenic plant; DNA replication; gene regulation; gene expression.

XX Myxococcus xanthus.

XX US6833447-B1.

FN 21-DEC-2004.

PD 10-JUL-2001; 2001US-00902540.

XX 10-JUL-2000; 2000US-0217883P.

PR (MONS) MONSANTO TECHNOLOGY LLC.

PA Goldman BS, Hinkle GT, Slater SC, Wiegand RC;

PI WPI; 2005-028716/03.

DR New substantially purified Myxococcus xanthus nucleic acid molecule
XX encoding a nitrite reductase, useful for determining gene expression,
PT identifying mutations in a gene of interest, and for constructing
PT mutations in a gene of interest.

XX Example 2; SEQ ID NO 10451; 25pp; English.

XX The invention relates to a substantially purified nucleic acid molecule
XX encoding a nitrite reductase of SEQ ID NO. 11926. Further disclosed is a
XX recombinant DNA construct for expression of a nitrite reductase gene in a
XX plant cell, and a plant cell comprising the recombinant DNA construct.CC The nucleic acid is useful for determining gene expression, identifying
CC mutations in a gene of interest, and for constructing mutations in a gene
CC of interest. Sequences given in records for SEQ IDs 9692-16825 represent
CC a group of 7134 Myxococcus xanthus proteins and peptides. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from USPTO
XX

SQ Sequence 226 AA;

Query Match 15.6%; Score 547.5; DB 9; Length 226;

Best Local Similarity 51.8%; Pred. No. 8,3e-43;
Matches 118; Conservative 30; Mismatches 73; Indels 7; Gaps 2;

QY 12 RSPRRVPRDRREGAVLAELAAVROPDVIYGLARGGLPVAMEVAAAALHAPLDAFVVR 71
DB 2 RSPRRVPRDRREGAVLAELAAVROPDVIYGLARGGLPVAMEVAAAALHAPLDAFVVR 71
QY 72 KLGAPEGHDEPAVGALASGGRVNVNDVVRGLRTTPOOLRDIAREGRELRRRESAYRGER 131
DB 72 KLGAPEGHDEPAVGALASGGRVNVNDVVRGLRTTPOOLRDIAREGRELRRRESAYRGER 131
QY 132 PPTDITGKTVIVDDGLATGASMFPAVQALRDAQPAQIVAVPAPESTCREPAGLVDDV 191
DB 132 PPTDITGKTVIVDDGLATGASMFPAVQALRDAQPAQIVAVPAPESTCREPAGLVDDV 191
QY 192 VCATMTPTPLAVGESFWDPROQVDEVRLLAFTPTGSPSRRAASTAA 239
DB 192 VCATMTPTPLAVGESFWDPROQVDEVRLLAFTPTGSPSRRAASTAA 239
QY 239 VCATMTPTPLAVGESFWDPROQVDEVRLLAFTPTGSPSRRAASTAA 239
DB 239 VCATMTPTPLAVGESFWDPROQVDEVRLLAFTPTGSPSRRAASTAA 239

RESULT 3
ID ADU49950 standard; protein; 217 AA.
XX
AC ADU49950;
XX
DT 27-JAN-2005 (first entry)
XX Methanobacterium barkeri PPPT-like protein.
XX

KW nematocide; purine/pyrimidine phosphoribosyl transferase-like;
KM PPPT-like polypeptide; nematode; therapy; enzyme.
OS Mehanosarcina barkeri.
XX
XX WO2004096835-A1.
PN
XX
XX 11-NOV-2004.
PD
XX
XX 31-MAR-2003; 2003WO-US009532.
PF
XX
XX 31-MAR-2003; 2003WO-US009532.
PR
XX
XX (DIVE-) DIVERGENCE INC.
PA
XX
XX Kloeck AP, Williams DJ, Salmon B,
PI
XX
XX WPI; 2005-012639/01.
DR
XX
XX Novel purified purine/pyrimidine phosphoribosyl transferase-like
PT polypeptide, useful for identifying compound altering activity of
PT purine/pyrimidine phosphoribosyl transferase-like polypeptide.
PS
XX Example; SEQ ID NO 16; 50pp; English.
XX
XX The invention describes a purified purine/pyrimidine phosphoribosyl
CC transferase (PPPT)-like polypeptide (I) comprising an amino acid sequence
CC being at least 80% identical to a fully defined sequence (S1) of 233, 233
CC or 229 amino acids as given in the specification. Also described are: an
CC isolated nucleic acid molecule (II) comprising a nucleotide sequence
CC encoding a polypeptide comprising any one of (S1); and an antibody (A1)
CC binding specifically to (I); a vector comprising (II); and a transgenic
CC cell or transgenic organism having in its genome a transgene containing
CC (II). (I) is useful for a method of measuring the binding of a test
CC compound to (I), which involves providing (I), contacting a test compound
CC to (I), and measuring the binding of the test compound to (I). The method
CC further involves measuring PPPT-like activity of (I). The method further
CC involves providing a second polypeptide, where the second polypeptide is
CC a plant or mammalian PPPT-like polypeptide, contacting the test compound
CC to the second polypeptide, and measuring the binding of the test compound
CC to the second polypeptide. (I) is useful for identifying a compound that
CC alters the activity of (I), which involves providing (I), contacting a
CC test compound to (I), and measuring a PPPT-like activity of (I), where a
CC change in PPPT-like activity relative to the PPPT-like activity of (I) in
CC the absence of the test compound is an indication that the test compound
CC alters the activity of (I). The method further involves providing a
CC second polypeptide, where the second polypeptide is a plant or mammalian
CC PPPT-like polypeptide, contacting the test compound to the second
CC polypeptide, and measuring PPPT-like activity of the second polypeptide.
CC A1 is useful as therapeutic compound to protect an animal from nematodes,
CC as reagents in experimental assays to detect presence of nematodes, as
CC tools to screen for expression of the gene product in nematodes, as a
CC purification tool of PPPT-like protein, and as PPPT inhibitors/activators
CC for therapeutic purposes. This is the amino acid sequence of
CC Mehanosarcina barkeri PPPT-like protein used in a comparison with
CC purine/pyrimidine phosphoribosyltransferase (PPPT) detailed in the
CC invention.
XX
XX
SQ Sequence 217 AA;
Query Match 14.7%; Score 514; DB 9; Length 217;
Best Local Similarity 47.2%; Pred. No. 1.2e-39;
Matches 101; Conservative 46; Mismatches 67; Indels 0; Gaps 0;

DB 123 SDSVTIILIDGLATGATMRAVAISAIKTKNPAKIVAAFTGAPDTCELFKMEVDEVIAT 182
OY 197 PTFPLAVGSFMDPRQVTDDEVRRLATPTAGS 230
DB 183 PEPFYGVAMYGNSQTTDEBEVCKLDKAPALRA 216
RESULT 4
AEB36491
ID AEB36491 standard; protein; 220 AA.
XX
XX AEB36491;
AC
XX
XX 08-SEP-2005 (first entry)
DT
XX
XX L. pneumonia protein SEQ ID NO 823.
DE
XX
XX detection; infection; Antibacterial; Vaccine.
KW
XX
XX Legionella pneumonia.
OS
XX
XX WO2005049642-A2.
PN
XX
XX 02-JUN-2005.
PD
XX
XX 23-SEP-2004; 2004WO-IB003578.
PF
XX
XX 21-NOV-2003; 2003FR-00013687.
PR
XX
XX (INSP) INST PASTEUR.
PA (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
PA (UPLY-) UNIV LYON 1 BERNARD CLAUDE.
PA (CNRS) CNRS CENT NAT RECH SCI.
XX
XX Buchrieser C, Tichit M, Etienne J, Ma L, Cazalec C, Glaeser P,
PI Ruenloek C, Bouchier C, Zidane N, Magnier A, Kunst F, Vandenesch F,
PI Jarraud S;
XX
XX WPI; 2005-388305/40.
DR
XX
XX New genome of Legionella pneumonia Paris strain and derived
PT polypeptides, useful for detection or identification of the strain and
PT for treatment and prevention of infections.
PS
XX Claim 3; SEQ ID NO 823; 660pp; English.
XX
XX The invention relates to an isolated or purified nucleotide sequences (I)
CC from Legionella pneumonia Paris strain. (I), and their related
CC sequences or fragments, are useful as primers and probes for detection
CC and amplification, including differentiation between the Paris and
CC Philadelphia strains of Legionella pneumonia and to prepare recombinant
CC (hybrid) polypeptides (II). (II) are also useful for preparation of
CC specific antibodies (Ab), also used for detection/identification of
CC Legionella, and some (I), specifically those involved in synthesis of
CC surface proteins, are targets for identification of inhibitors. (II), or
CC vectors that contain (I), are useful as vaccines and immunogenic
CC compositions, for treatment and prevention of infections by L.
CC pneumonia. The present sequence represents the amino acid sequence of a
CC L. pneumonia protein.
XX
XX
SQ Sequence 220 AA;
Query Match 13.2%; Score 464; DB 9; Length 220;
Best Local Similarity 45.8%; Pred. No. 6.4e-35;
Matches 98; Conservative 40; Mismatches 74; Indels 2; Gaps 1;

```
Db          59  IVRKLVPGHEBLAFGALAIKIKVYNEBLLQYIQLSOKDINOVLQSOENELARRNAYR 118
Qy          129  GERPPDITGKTVIVDDIGLATGASMPAAVQALRDAQPAQIVIAVPAPESTCFEAGLV 188
           119  GDRKFPDLKNKIVILVDDIGLATGASMPAAVQALRDAQPAQIVIAVPAPESTCFEAGLV 178
Qy          189  DDVVCATMPPTFLAVGESFMDPROVTDDEVRRL 222
           179  DEVICLSPMSFNAVGLWYEDFSQTEDEEVLYLL 212
Db

RESULT 5
AEB39892
ID  AEB39892 standard; protein; 216 AA.
XX
AC  AEB39892;
XX
DT  08-SEP-2005 (first entry)
XX
DE  L. pneumophila protein SEQ ID NO 4224.
XX
KM  detection; infection; Antibacterial; Vaccine.
XX
OS  Legionella pneumophila.
XX
PN  W02005049642-A2.
XX
PD  02-JUN-2005.
XX
PF  23-SEP-2004; 2004WO-IB003578.
XX
PR  21-NOV-2003; 2003FR-00013687.
XX
PA  (INSP ) INST PASTEUR.
XX
PA  (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.
XX
PA  (IUYL-) UNIV LYON 1 BERNARD CLAUDE.
XX
PA  (CNRS ) CNRS CENT NAT RECH SCT.
XX
PI  Buchrieser C, Tichit M, Etienne J, Ma L, Cazalet C, Glaser P,
PI  Rusniok C, Bouchier C, Zidane N, Magnier A, Kunst F, Vandenesch F,
PI  Jarrard S,
XX
DR  WPI; 2005-388305/40.
XX
XX
PT  New genome of Legionella pneumophila Paris strain and derived
PT  polypeptides, useful for detection or identification of the strain and
PT  for treatment and prevention of infections.
XX
XX
PS  Claim 3; SEQ ID NO 4224; 660bp; English.
XX
XX
CC  The invention relates to an isolated or purified nucleotide sequences (I)
CC  from Legionella pneumophila Paris strain. (I), and their related
CC  sequences or fragments, are useful as primers and probes for detection
CC  and amplification, including differentiation between the Paris and
CC  Philadelphia strains of Legionella pneumophila and to prepare recombinant
CC  (hybrid) polypeptides (II). (II) are also useful for preparation of
CC  specific antibodies (Ab), also used for detection/identification of
CC  Legionella, and some (I), specifically those involved in synthesis of
CC  surface proteins, are targets for identification of inhibitors. (II), or
CC  vectors that contain (I), are useful as vaccines and immunogenic
CC  compositions, for treatment and prevention of infections by L.
CC  pneumophila. The present sequence represents the amino acid sequence of a
CC  L. pneumophila protein.
XX
XX
SQ  Sequence 216 AA:
Query Match 13.2%; Score 463; DB 9; Length 216;
Best Local Similarity 46.8%; Pred. No. 7.7e-35;
Matches 96; Conservative 37; Mismatches 72; Indels 0; Gaps 0;
Qy          18  FQDRREGRVYAEELAAVRDQPDYIVIGIARGSLPVAMEVAAALHAPLDAFVVRKLGAPG 77
           4  FQDRREGRVYAEELAAVRDQPDYIVIGIARGSLPVAMEVAAALHAPLDAFVVRKLGAPG 63
Db
```

```
Qy          78  HDEFAVAGALASGRGVVNDVVRGRLITPQQLRIABREBERELLRESAARGSRPPDIT 137
           64  HEBIARFGALAIKIKVYNEBLLQYIQLSOKDINOVLQSOENELARRNAYRGGKPEPDLK 123
Db
Qy          138  GKTIVIVDDIGLATGASMPAAVQALRDAQPAQIVIAVPAPESTCFEAGLVVCAATMP 197
           124  NKIVILVDDIGLATGASMPAAVQALRDAQPAQIVIAVPAPESTCFEAGLVVCAATMP 183
Db
Qy          198  TPFLAVGESFMDPROVTDDEVRRL 222
           184  MSFNAVGLWYEDFSQTEDEEVLYLL 208
Db

RESULT 6
AEB95281
ID  AEB95281 standard; protein; 443 AA.
XX
AC  AEB95281;
XX
DT  06-OCT-2005 (first entry)
XX
DE  M tuberculosis purine/pyrimidine phosphoribosyl transferase - SEQ ID 10.
XX
KM  purine/pyrimidine phosphoribosyl transferase; vaccine; pesticide;
XX  nematode infection; nematocide; enzyme.
XX
OS  Mycobacterium tuberculosis.
XX
PN  US2005172363-A1.
XX
PD  04-AUG-2005.
XX
PF  16-FEB-2005; 2005US-00058869.
XX
PR  30-MAR-2001; 2001US-0280192P.
XX
PR  29-MAR-2002; 2002US-00113201.
XX
PR  21-JAN-2003; 2003US-00348317.
XX
PA  (DIVE-) DIVERGENCE INC.
XX
PI  Kloeck AP, Williams DJ, Salmon B;
XX
DR  WPI; 2005-532486/54.
XX
XX
PT  New isolated purine/pyrimidine phosphoribosyl transferase nucleic acid
PT  and polypeptide, useful as vaccines, pesticides for plants, soil, or
PT  seeds to confer nematode resistance, or drugs treating diseases caused by
PT  nematodes.
XX
XX
PS  Example; SEQ ID NO 10; 21bp; English.
XX
XX
CC  The invention comprises the amino acid and coding sequence of a
CC  purine/pyrimidine phosphoribosyl transferase. The DNA and protein
CC  sequences of the invention are useful as vaccines, pesticides, and drugs
CC  for protecting plants and animals against nematode infection. The present
CC  amino acid sequence represents a Mycobacterium tuberculosis
CC  purine/pyrimidine phosphoribosyl transferase enzyme.
XX
XX
SQ  Sequence 443 AA:
Query Match 12.3%; Score 431; DB 9; Length 443;
Best Local Similarity 32.4%; Pred. No. 2.6e-31;
Matches 129; Conservative 52; Mismatches 161; Indels 56; Gaps 11;
Qy          16  RVFDRREGRVYAEELAAVRDQPDYIVIGIARGSLPVAMEVAAALHAPLDAFVVRKLG 75
           2  KLPDDGDMGRQKLAQLSGSA-VVVGLEPRGVPVAFVAKSLQAPLDVLVVRKLG 60
Db
Qy          76  PGHDEFAVAGALASGRGVVNDVVRGRLITPQQLRIABREBERELLRESAARGSRPPD 135
           61  PGHDEFAVAGALASGRGVVNDVVRGRLITPQQLRIABREBERELLRESAARGSRPPD 120
Db
```

QY 136 ITGKTVIVDDGLATGASMPFAVQALRDAPQAVIAVPAAPSTCFEAGLVDDVVCAT 195
 Db 121 LTGRIAVTVDDGATGATAKACQVABAGADKVLAVPIGPDDIVARFAGYADDEVCLA 180
 QY 196 MPTFLAVGESFMDFRQVTEDEVRRLATPTAGSLRRPAASTADVLRRVAIDAPGV 255
 Db 181 TPALFFAVGGYRNFTQTSDEVAFL-----DRAHRDPAEAGALIDAAADPP 227
 QY 256 THEVLAEVGDARIVLIGESSHGTHEFYQARAAMTOWLIEKGFQAVA-AEADMPDAYRV 314
 Db 228 LRDEEVGVVAGP-----VPVAGHLT-----VPEKPRGIVVFAGSGSSSRHSI 269
 QY 315 -NRVY-----RGLGEDTNDALSGFER-----PPAMMRNTVVRDPEVRLRTNQRYES 363
 Db 270 RNRVVAEVLTGAGFATLLFDLLTPEEERNRANVFDELLASRLI-DVTGMLATQ----PD 324
 QY 364 GALRQAGFYGLDLYSLHRSIQEVI SYLDKVPRAAARA 401
 Db 325 TALSIPVGIFG-----ASTGAGALVAAADPRVNVRA 355

RESULT 7

AE93950 ID AEB93950 standard; protein; 443 AA.

AC AEB93950;

DT 06-OCT-2005 (first entry)

DE Mycobacterium tuberculosis PPPT-like protein, SEQ ID NO: 10.

KW Screening; antibody production; nematode infection; nematocide;

KM infection; purine/pyrimidine phosphoribosyl transferase-like protein.

OS Mycobacterium tuberculosis.

PN US2005172350-A1.

PD 04-AUG-2005.

PF 27-JAN-2005; 2005US-00045942.

PR 30-MAR-2001; 2001US-0280192P.

PR 29-MAR-2002; 2002US-00113201.

PR 21-JAN-2003; 2003US-00347776.

PA (DIVE-) DIVERGENCE INC.

PI Kloeck AP, Williams DJ, Salmon B;

DR WPI; 2005-563091/57.

PT Screening for inhibitors of nematode purine/pyrimidine phosphoribosyl transferase (PPPT), useful for treating nematode infections (especially in plants), comprises contacting a test compound with the enzyme from *Meloidogyne incognita*.

PS Example; SEQ ID NO 10; 29PP; English.

CC The invention relates to a method of screening for inhibitors of nematode purine/pyrimidine phosphoribosyl transferase (PPPT, also known as PRTase) which involves contacting a test compound with the enzyme from *Meloidogyne incognita*. The invention also relates to methods for antibody production. The method is useful for screening for compounds that alter the activity of PPPT. Compounds identified by the method, especially compounds that have a more powerful effect on the nematode enzyme than on plant or mammalian PPPTs, are useful for treating nematode infections, especially in plants. The present sequence is the *Mycobacterium tuberculosis* PPPT-like protein.

XX Sequence 443 AA;

Query Match 12.3%; Score 431; DB 9; Length 443;

Best Local Similarity 32.4%; Pred. No. 2, 6e-31;
 Matches 129; Conservative 52; Mismatches 161; Indels 56; Gaps 11;

QY 16 RVFRDRREAGRVLAELLAAYRDQPDVVLIGIABGGLPVAVEVAALHAPIDAFVVRXIGA 75
 Db 2 KLPDDRGRDAGRQLAQRLAQLSGKA-VVVLGLPRGGVVAEVAKSLQAPLDVLVVRXLGV 60
 QY 76 PGHDEFVAGLAAAGGRVNVDDVYRGLRITPQQLRDIAREGBELLRRESAYGGERPTD 135
 Db 61 PFOPELAFGAIGBDGVRVLNDVVRGTHLDAAMDAVERQQLTELQRRARFRGRDRIP 120
 QY 136 ITGKTVIVDDGLATGASMPFAVQALRDAPQAVIAVPAAPSTCFEAGLVDDVVCAT 195
 Db 121 LTGRIAVTVDDGATGATAKACQVABAGADKVLAVPIGPDDIVARFAGYADDEVCLA 180
 QY 196 MPTFLAVGESFMDFRQVTEDEVRRLATPTAGSLRRPAASTADVLRRVAIDAPGV 255
 Db 181 TPALFFAVGGYRNFTQTSDEVAFL-----DRAHRDPAEAGALIDAAADPP 227
 QY 256 THEVLAEVGDARIVLIGESSHGTHEFYQARAAMTOWLIEKGFQAVA-AEADMPDAYRV 314
 Db 228 LRDEEVGVVAGP-----VPVAGHLT-----VPEKPRGIVVFAGSGSSSRHSI 269
 QY 315 -NRVY-----RGLGEDTNDALSGFER-----PPAMMRNTVVRDPEVRLRTNQRYES 363
 Db 270 RNRVVAEVLTGAGFATLLFDLLTPEEERNRANVFDELLASRLI-DVTGMLATQ----PD 324
 QY 364 GALRQAGFYGLDLYSLHRSIQEVI SYLDKVPRAAARA 401
 Db 325 TALSIPVGIFG-----ASTGAGALVAAADPRVNVRA 355

RESULT 8

ADU49942 ID ADU49942 standard; protein; 233 AA.

AC ADU49942;

DT 27-JAN-2005 (first entry)

DE *Meloidogyne javanica* purine/pyrimidine phosphoribosyl transferase.

KW nematocide; purine/pyrimidine phosphoribosyl transferase-like;

KM PPPT-like polypeptide; nematode; therapy; enzyme.

OS *Meloidogyne javanica*.

PN WO2004096835-A1.

PD 11-NOV-2004.

PF 31-MAR-2003; 2003WO-US009532.

PR 31-MAR-2003; 2003WO-US009532.

PA (DIVE-) DIVERGENCE INC.

PI Kloeck AP, Williams DJ, Salmon B;

DR WPI; 2005-012639/01.

DR N-PSDB; ADU49939, ADU49954.

PT Novel purified purine/pyrimidine phosphoribosyl transferase-like polypeptide, useful for identifying compound altering activity of purine/pyrimidine phosphoribosyl transferase-like polypeptide.

PS Claim 4; SEQ ID NO 5; 50PP; English.

CC The invention describes a purified purine/pyrimidine phosphoribosyl transferase (PPPT-like polypeptide (I) comprising an amino acid sequence being at least 80% identical to a fully defined sequence (S1) of 233, 233 or 229 amino acids as given in the specification. Also described are: an isolated nucleic acid molecule (II) comprising a nucleotide sequence

XX WPI; 2005-563091/57.
DR P-PSDB; AEB93955.
XX Screening for inhibitors of nematode purine/pyrimidine phosphoribosyl
PT transferase (PPT), useful for treating nematode infections (especially
PT in plants), comprises contacting a test compound with the enzyme from
PT Meloidogyne incognita.
XX
XX Example; SEQ ID NO 5; 29pp; English.
XX The invention relates to a method of screening for inhibitors of nematode
CC purine/pyrimidine phosphoribosyl transferase (PPT, also known as PRTase)
CC which involves contacting a test compound with the enzyme from
CC Meloidogyne incognita. The invention also relates to methods for antibody
CC production. The method is useful for screening for compounds that alter
CC the activity of PPT. Compounds identified by the method, especially
CC compounds that have a more powerful effect on the nematode enzyme than on
CC plant or mammalian PPTs, are useful for treating nematode infections,
CC especially in plants. The present sequence is the root-knot nematode PPT
CC -like protein. Note: The present sequence is the SEQ ID NO: 5 which is
CC shown on figure 2. This sequence differs from the SEQ ID NO: 5 given in
CC the sequence listing (see AEB93945).
XX
SQ Sequence 233 AA;
Query Match 12.0%; Score 420.5; DB 9; Length 233;
Best Local Similarity 42.4%; Pred. No. 9.3e-31;
Matches 89; Conservative 43; Mismatches 75; Indels 3; Gaps 2;
QY 12 RSPRRVFRDRAGRLVLAELIAAYRDPD-VITGLARGGLPVAMEVAALHAPLDAFV 70
DB 5 RAAAPAFKDRHDAQGLAEALKNFRSQDRKVVVALPRGVPVAFEVAKSGAPLDLIMV 64
QY 71 RKLAPGDEFAVAGLASG--RVVNDVVRGLRITPQQLRDLAEBRGRELLRESAYR 128
DB 65 RKIGPGEHEBYGIGAVNEGPNPELVNEDAVKVTOPPEGIVQAMMEKOLKEIARORNYL 124
QY 129 GERPTDTTGKTVIVDDGLATGASMPAAVOALRDAPQIVIAVPAPESTCEFAGLV 188
DB 125 GDRPPLSLAGKIAIVDDGATGTAARVAMKALRKXNAKLLASPLAPSTLAELEAREG 184
QY 189 DDVVCATMPTPLAVGESFMDPRQVTDEEV 218
DB 185 NEVLVLETPNFSAVGLHYTKFDQTSDBEV 214
RESULT 11
ABM96175
ID ABM96175 standard; protein; 229 AA.
XX
XX ABM96175;
XX
XX 02-JUN-2005 (first entry)
XX
XX M. xanthus protein sequence, seq id 15374.
XX
XX Transgenic plant; DNA replication; gene regulation; gene expression.
XX
XX Myxococcus xanthus.
XX
XX US683447-B1.
XX
XX 21-DEC-2004.
XX
XX 10-JUL-2001; 2001US-00902540.
XX
XX 10-JUL-2000; 2000US-0217883P.
XX
XX (MONS) MONSANTO TECHNOLOGY LLC.
XX
XX Goldman BS, Hinkle GJ, Slater SC, Wiegand RC;
XX

DR WPI; 2005-028716/03.
XX
XX New substantially purified Myxococcus xanthus nucleic acid molecule
PT encoding a nitrite reductase, useful for determining gene expression,
PT identifying mutations in a gene of interest, and for constructing
PT mutations in a gene of interest.
XX
XX Example 2; SEQ ID NO 15374; 25pp; English.
XX
XX The invention relates to a substantially purified nucleic acid molecule
CC encoding a nitrite reductase of SEQ ID NO. 11926. Further disclosed is a
CC recombinant DNA construct for expression of a nitrite reductase gene in a
CC plant cell, and a plant cell comprising the recombinant DNA construct.
CC The nucleic acid is useful for determining gene expression, identifying
CC mutations in a gene of interest, and for constructing mutations in a gene
CC of interest. Sequences given in records for SEQ IDs 9692-16825 represent
CC a group of 7134 Myxococcus xanthus proteins and peptides. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from USPTO
XX
SQ Sequence 229 AA;
Query Match 11.9%; Score 417.5; DB 9; Length 229;
Best Local Similarity 42.5%; Pred. No. 1.7e-30;
Matches 96; Conservative 39; Mismatches 86; Indels 5; Gaps 3;
QY 18 FRDRERAGRLVLAELI--AAVRDQPDVITGLARGGLPVAMEVAALHAPLDAFVRYKGA 75
DB 3 FEDRVDAAGRLAQRLARRRGVTGE-GLTVLALPRGVPVAFEVATLALGAPLDVWVRKGV 61
QY 76 PGHDEFVAGLASGGRVNVNDVVRGLRITPQQLRDLAEBRGRELLRESAYRREPERTD 135
DB 62 PGYELDGLGAVBEGVAFVRRRLMDEVGVTEEDMQGLVRKQTBKRVAFRFGIEAPR 121
QY 136 ITGKTVIVDDGLATGASMPAAVOALRDAPQIVIAVPAPESTCEFAGLVDVVCAT 195
DB 122 IEGQRITLVDDGATGTAARVAMKALRKXNAKLLASPLAPSTLAELEAREG 181
QY 196 MPTPLAVGESFMDPRQVTDEEVRLIATP--TAGSLRRPAASTA 239
DB 182 STPLVAIGQVADFPQVPEAVATLLAQRLTLGHRHAPPGDTA 227
RESULT 12
ADU49941
ID ADU49941 standard; protein; 233 AA.
XX
XX ADU49941;
XX
XX 27-JAN-2005 (first entry)
XX
XX Meloidogyne incognita purine/pyrimidine phosphoribosyltransferase.
XX
XX nematocide; purine/pyrimidine phosphoribosyl transferase-like;
XX PPT-like polypeptide; nematode; therapy; enzyme.
XX
XX Meloidogyne incognita.
XX
XX WO2004096835-A1.
XX
XX 11-NOV-2004.
XX
XX 31-MAR-2003; 2003WO-US009532.
XX
XX 31-MAR-2003; 2003WO-US009532.
XX
XX (DIVE-) DIVERGENCE INC.
XX
XX Kloeck AP, Williams DJ, Salmon B;
XX
XX WPI; 2005-012639/01.
XX
XX N-PSDB; ADU49938, ADU49953.
XX

PT Novel purified purine/pyrimidine phosphoribosyl transferase-like polypeptide, useful for identifying compound altering activity of purine/pyrimidine phosphoribosyl transferase-like polypeptide.

PS Claim 4; SEQ ID NO 4; 50pp; English.

The invention describes a purified putine/pyridine phosphoribosyl transferase (PPPT)-like polypeptide (I) comprising an amino acid sequence being at least 80% identical to a fully defined sequence (S1) of 233, 233 or 235 amino acids as given in the specification. Also described are: an isolated nucleic acid molecule (II) comprising a nucleotide sequence encoding a polypeptide comprising any one of (S1); and an antibody (A1) binding specifically to (I); a vector comprising (II); and a transgenic cell or transgenic organism having in its genome a transgene containing (II). (I) is useful for a method of measuring the binding of a test compound to (I), which involves providing (I), contacting a test compound to (I), and measuring the binding of the test compound to (I). The method further involves measuring PPPT-like activity of (I). The method further involves providing a second polypeptide, where the second polypeptide is a plant or mammalian PPPT-like polypeptide, contacting the test compound to the second polypeptide, and measuring the binding of the test compound to the second polypeptide. (I) is useful for identifying a compound that alters the activity of (I), which involves providing (I), contacting a test compound to (I), and measuring a PPPT-like activity of (I), where a change in PPPT-like activity relative to the PPPT-like activity of (I) in the absence of the test compound is an indication that the test compound alters the activity of (I). The method further involves providing a second polypeptide, where the second polypeptide is a plant or mammalian PPPT-like polypeptide, contacting the test compound to the second polypeptide, and measuring PPPT-like activity of the second polypeptide. A1 is useful as therapeutic compound to protect an animal from nematodes, as reagents in experimental assays to detect presence of nematodes, as tools to screen for expression of the gene product in nematodes, as a purification tool of PPPT-like protein, and as PPPT inhibitors/activators for therapeutic purposes. This sequence encodes is the amino acid sequence of Meloidogyne incognita putine/pyrimidine phosphoribosyltransferase (PPPT).

SQ Sequence 233 AA:

Query Match	11.9%	Score 417.5	DB 9	Length 233
Best Local Similarity	41.9%	Pred. No. 1.8e-30		
Matches 88	Conservative 44	Mismatches 75	Indels 3	Gaps 2

QY 12 RSRRFRKRRRNGRGLTAEILLAAYKDPD -YIVGLAGGLPVAMEVAAAALHAPLDLDFV 70
Db 5 RAATAPFKKRHDGQGLTALNKNFQSRKRVVLLPFGGVPAVFEAKSLGAPLDLLMV 64
QY 71 RKTAGPGBDFEAFGALASG--RVVNDVDVNGRLITPQGLDRLAERREGLREBSAYR 128
Db 65 RKTAGGHEEYTGGAIVVEGNPELVNNEAVKTYTPPGGYQAMMEKOLKEIARQRRYL 124
QY 129 GERPRDTIGKTVIVDDGLATGASMFVAQALRDAPQIVIAVPAAPBSTCEFAGLV 168
Db 125 GDRPPLSLAKIAIIVDDGIATGTFARVAMKRLRQNKVAKALLSLAPSDTLAEIARAG 184
QY 189 DDVVCATMPPLPAVBESFMDPROUTDEEV 218
Db 185 NEVLVLETPEPNSFAGVLAHTKDDQDSDEV 214

RESULT 13
AEB95271
ID AEB95271 standard; protein; 233 AA

DT 06-OCT-2005 (first entry)

DE M incognita purine/pyrimidine phosphoribosyl transferase - SEQ ID 4.

purine/pyrimidine phosphoribosyl transferase; vaccine; pesticide; nematode infection; nematocide; enzyme.

XX Meloidogyne incognita.
OS

PN US2005172363-A1.

PD 04-AUG-2005

PF 16-FEB-2005; 2005US-00058869

PR 30-MAR-2001; 2001US-0280192P

PR 21-JAN-2003; 2003US-00348317.

PA (DIVE-) DIVERGENCE INC.

PI Klock AP, Williams DJ,

DR WPI; 2005-532486/54.

XX

PT and polypeptide

PT nematodes

PS Claim 5; SEQ ID NO 4; 21pp; English.

CC The invention comprises the amino acid and coding sequence of a

CC sequences of the invention are useful as vaccines, pesticides, and drugs.

CC amino acid sequence represents the Meloidogyne incognita

XX

SQ Sequence 233 AA;

Query Match	11.9%	Score 417.5	DB 9	Length 233
Best Local Similarity	41.9%	Pred. No. 1.8e-30		
Matches 88	Conservative 44	Mismatches 75	Indels 3	Gaps 2

[illegible]

RESULT 14
AEB93944
ID AEB93944 standard; protein; 233 AA

DT 06-OCT-2005 (first entry)

DE Southern root-knot nematode PPPT-like protein, SEQ ID NO: 4.

KW Screening; antibody production; nematode infection; nematocide;

OS Meloidogyne incognita.

PN US2005172350-A1.

Query	Match	11.9%	Score 417.5	DB 9	Length 233
Best Local Similarity	41.9%	Pred. No. 1.8e-30			
Matches 88	Conservative 44	Mismatches 75	Indels 3	Gaps 2	
QY	12	RSPPRVFDRRBAGVLAELLAAYRDPD-VVIGLGLSGLPVAMEVAAALAPDADFV	70		
DB	5	RAAARAPFDRDRDAAGKLAELKNFSSQDKRVVLLPRGSPVAVAEVAKSLGAPDILMV	64		
QY	71	RKLGA PGHDEPAVGLASGG--RVVVNDVYKGLRITPQULDIERGGRELLRESAYR	128		
DB	65	RKIGAPGHEEVIGIGAVVEGNPELVNMDAVKYTPPGSYVQAMMEKQULKEIARQRNYL	124		
QY	129	GERPEPTDTGKTIVYVDDGLATGASMPAAVQALRPAQIVIAVPAAPSESTCFEAGLV	188		
DB	125	GDRPEPLSLAGIKAIYVDDGIATGCTARVAMKRLRQKNYAKALLASPLAPSDTLAEIARAEG	184		
QY	189	DDVVCATMPETPELVAVGSEFMDPROVTDREV	218		
DB	185	NEVLVLETPNPSAVGLHYTFQDTSDEV	214		
RESULT 15					
ID	ADB74647	standard; protein; 289 AA.			
XX	ADB74647				
XX	ADB74647				
DT	04-DEC-2003	(first entry)			
DB	Mycobacterium leprae	non-naturally occurring peptide #347.			
XX					
KW	Non-naturally occurring peptide; anion pump protein; tuberculosis;				
XX	hypoensitivy reaction; tuberculostatic.				
OS	Mycobacterium leprae.				
PN	US6583266-B1				

XX 24-JUN-2003.
PD
XX 16-SEP-1994; 94US-0011731.
XX
PF
XX 19-AUG-1993; 93US-00109181.
PR
XX 22-OCT-1993; 93US-00142558.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
F1 Smith DR, Mao J;
XX
XX WPI; 2003-656441/62.
XX
PT New Mycobacterium tuberculosis anion pump peptide useful for as
PT tuberculosis vaccine and diagnosis of tuberculosis infection.
XX
PS
XX Disclosure; SEQ ID NO 396; 26pp; English.
XX
CC The invention relates to a non-naturally occurring peptide of
CC Mycobacterium tuberculosis comprising an amino acid sequence
CC corresponding to an anion pump protein. The invention also relates to a
CC non-naturally occurring nucleic acid corresponding to a DNA sequence of
CC Mycobacterium tuberculosis or Mycobacterium leprae. The new peptide is
CC useful as a vaccine against Mycobacterium tuberculosis or Mycobacterium
CC leprae or for screening for new tuberculosis drugs. Purified proteins
CC derived from the sequences of the invention may elicit a specific immune
CC response. The peptide may also be used to detect hypersensitivity
CC reactions of individuals exposed to Mycobacterium tuberculosis or
CC Mycobacterium leprae. The proteins and peptides may be affixed to solid
CC supports to detect antibodies typical of hypersensitivity reactions, from
CC a patient's sera. This sequence represents Mycobacterium leprae non-
CC naturally occurring peptide of the invention. Note: The sequence data for
CC this patent did not form part of the printed specification but was
CC obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html.

```

XX      SQ      Sequence 289 AA;
Query Match      11.8%; Score 415.5; DB 7; Length 289;
Best Local Similarity 41.6%; Pred. No. 3.9e-30;
Matches 97; Conservative 33; Mismatches 90; Indels 13; Gaps 3

QY      12  RSPRRVRRDRREAGRVLAELLAAYRDOPDIVLGLARGLEFVAMEVAALAHAPLDAFYVR 71
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      53  RNSRRLFDPRDARDAQRHLAERLESIRGK-DIVVLGLPGGVFAVEFAKALRAPLDIVLR 111
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY      72  KLGAPGHDEFVAGLLAGGRVYVNDVYVGRGRIITPQQLRDIAEERGRLLRESAYRBER 131
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      112  KLGAPFOPELAFGAIGREGVAVTDSVVAEADLCODEAALETORALIMRSERFHRRH 171
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY      132  PPTDITGKTVVNDGGLATGASMAFAVOALDADPAQIVIAVPAAPESTCRBFAGLVNDV 191
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      172  DRAITGRIAVIVDDGINTGATTAASCQVAPAGVSKVLAVPITGGRITFARFAGYADEV 231
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY      192  VCATMPPTPLFVAGSESFWDFRQVTDEBVRLL-----ATPTAG-PSLR 232
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      232  VCLHTPAPFCVAGGVCNFTQTSDAEVIALLDRAREGSESATYATLGDPOIR 284
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Search completed: March 23, 2006, 05:10:44
Job time : 183.055 secs

```

This page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: March 23, 2006, 05:11:18 ; Search time 27.0938 Seconds
(without alignments)
2418.401 Million cell updates/sec

Title: US-10-617-038-23

Perfect score: 3508

Sequence: 1 VLMTPAADVTRSPRRVFRD.....LEVTSRWIAGENPPTPTGL 681

Scoring table: BLOSUM62

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_80:*

1: p1r1:*

2: p1r2:*

3: p1r3:*

4: p1r4:*

SUMMARIES

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match Length	ID	Description
1	3505	99.9	681 2	E70942 hypothetical prote
2	894.5	25.5	658 2	H95852 hypothetical fusio
3	780.5	22.2	554 2	T748790 hypothetical prote
4	668	19.0	352 2	C70578 hypothetical prote
5	540	15.4	221 2	A62055 hypothetical prote
6	431	12.3	443 2	B70933 hypothetical prote
7	276	7.9	554 2	A75296 probable erythromy
8	241	6.9	434 2	F84332 succinoglycan bios
9	224	6.4	419 1	B5E6RM erythromycin ester
10	220	6.3	175 2	D70305 conserved hypotnet
11	199	5.7	396 2	H67356 hypothetical prote
12	185	5.3	383 2	E75589 hypothetical prote
13	178.5	5.1	387 2	T36697 hypothetical prote
14	159	4.5	446 2	A69750 erythromycin ester
15	157	4.5	240 2	F64671 conserved hypotnet
16	150	4.3	234 2	F71845 hypothetical prote
17	145.5	4.1	621 2	T46851 conserved hypotnet
18	140	4.0	919 2	T37062 probable transcrip
19	136	3.9	574 2	H83117 probable ATP-bindi
20	133.5	3.8	1238 2	T03465 probable exonuclea
21	132	3.8	1185 2	T36761 probable DNA polym
22	131	3.7	1275 2	D38164 cobN protein - pse
23	130.5	3.7	8563 2	T30226 polypeptide synthas
24	129.5	3.7	7576 2	T17428 FK506 polypeptide s
25	129	3.7	423 2	T11214 3'-oxocyl-lacyl-ca
26	128.5	3.7	537 2	T28683 hypothetical prote
27	128.5	3.7	749 2	A55953 probable bifunctio
28	128.5	3.7	885 2	A64373 leucine-CRNA synth
29	128.5	3.7	3413 2	T17467 rifamycin polypekti

30	128	3.6	7463 2	T36248 CDA peptide synthe
31	127.5	3.6	3164 1	WMBEH6 U36 protein - hum
32	127	3.6	384 2	AG2790 conserved hypotnet
33	127	3.6	384 2	F97569 hypothetical prote
34	127	3.6	1361 2	T29435 hypothetical prote
35	127	3.6	1632 2	C70752 probable ctpi prot
36	127	3.6	1937 2	T03224 probable polyketid
37	126.5	3.6	3739 2	T17410 polyketide synthas
38	126	3.6	572 2	T36883 probable DNA repai
39	125.5	3.6	682 2	T36877 probable regulator
40	125	3.6	1116 2	AC2921 cobalamin biosynth
41	125	3.6	1116 2	D97695 cobN protein homol
42	125	3.6	1321 2	T10929 3C3.20c protein -
43	125	3.6	2117 2	T36180 CDA peptide synthe
44	124.5	3.5	6420 2	T30283 polypeptide synthas
45	124	3.5	502 2	T50754 protochlorophyllid

ALIGNMENTS

RESULT 1				
E70942	hypothetical protein Rv2030c - Mycobacterium tuberculosis (strain H37Rv)			
C:Species:	Mycobacterium tuberculosis			
C:Date:	17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 31-Dec-2004			
C:Accession:	E70942			
R:Colo, S.T.: Broesch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holtroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Nature 393, 537-544, 1998				
A:Authors:	Squares, R.; Suleston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.			
A:Title:	Deciphering the biology of Mycobacterium tuberculosis from the complete genome			
A:Reference number:	A70500; PMID:98295987; PMID:9634230			
A:Accession:	E70942			
A:Status:	preliminary; nucleic acid sequence not shown; translation not shown			
A:Molecule type:	DNA			
A:Residues:	1-681 <COL>			
A:Cross-references:	UNIPROT:O53475; UNIPARC:UPI00000D108D; GB:AL021899; GB:AL123456; NID			
A:Experimental source:	strain H37Rv			
C:Genetic:				
A:Gene:	Rv2030c			
C:Superfamily:	uncharacterized conserved protein with erythromycin esterase domain			
Query Match 99.9%; Score 3505; DB 2; Length 681;				
Best Local Similarity 99.9%; Pred. No. 1.4e-217;				
Matches 680; Conservative 1; Mismatches 0; Indels 0; Gaps 0;				
QY	1	VLMTPAADVTRSPRRVFRDREAGRVLAELAAVRDQPVIVGLARGGLPVAMEVAAA	60	
DB	1	MLMTAAADVTRSPRRVFRDREAGRVLAELAAVRDQPVIVGLARGGLPVAMEVAAA	60	
QY	61	LHAPLDAFVVRKUGAPGHDFPAVGALASGGRVVYNDVVRGRTTPQQLDIAREBEREL	120	
DB	61	LHAPLDAFVVRKUGAPGHDFPAVGALASGGRVVYNDVVRGRTTPQQLDIAREBEREL	120	
QY	121	LRRSSAYRGRPPPTIDIGKTVIVVDGLATGASFAAVALRDPQPIYVLAAPAEEST	180	
DB	121	LRRSSAYRGRPPPTIDIGKTVIVVDGLATGASFAAVALRDPQPIYVLAAPAEEST	180	
QY	181	CRFPAGLVDDVVCATMPTPLAVGESFMDRQVYVDEVRRLATPTGASLRPAASTAA	240	
DB	181	CRFPAGLVDDVVCATMPTPLAVGESFMDRQVYVDEVRRLATPTGASLRPAASTAA	240	
QY	241	DVLRRAVDAIPGGVPTHEVLAELVGDARIVLIGESSHGTHFYOARAAATQWLEEGFG	300	
DB	241	DVLRRAVDAIPGGVPTHEVLAELVGDARIVLIGESSHGTHFYOARAAATQWLEEGFG	300	
QY	301	AVAAEADMPDAPYRNRVVRGIGEDTNADEALSGPERPAMMMRTVVRDPEVMLRTNOR	360	
DB	301	AVAAEADMPDAPYRNRVVRGIGEDTNADEALSGPERPAMMMRTVVRDPEVMLRTNOR	360	
QY	361	YESGALRQAGFYGLDLYSLHRSIOEIVSYLDKVPRAAARARAYACFDHACDGAAYG	420	

```

Db      361 YESALGKAGFYGLDLSLHRSIOEVLSTLDKVDPRAPARAPACFDHACADGGQAYG 420
Qy      421 FAAAFGAGPSCEREAVEQLVDVORNALAYARODGLADELFYAQONAAQTVRDAEYVYRA 480
Db      421 FAAAFGAGPSCEREAVEQLVDVORNALAYARODGLADELFYAQONAAQTVRDAEYVYRA 480
Qy      481 MFSGRVTSWNLRDQMAQTGLSLTLHLDRHLDPAPARIVVAHNSHVGDAATEVMADGQ 540
Db      481 MFSGRVTSWNLRDQMAQTGLSLTLHLDRHLDPAPARIVVAHNSHVGDAATEVMADGQ 540
Qy      541 LTTGQIVRERYGDSRSIGFSTYGTATYAASEMGIAQRKAVRPALHSGVVELFHQTADS 600
Db      541 LTTGQIVRERYGDSRSIGFSTYGTATYAASEMGIAQRKAVRPALHSGVVELFHQTADS 600
Qy      601 FLVARSRLDAEAPLDVRLGRAIGVYLPATERQSHLHVRAPADQFPMHIDQTRALE 660
Db      601 FLVARSRLDAEAPLDVRLGRAIGVYLPATERQSHLHVRAPADQFPMHIDQTRALE 660
Qy      661 PLEVTSRWIAGENPEYPTGL 681
Db      661 PLEVTSRWIAGENPEYPTGL 681

```

RESULT 2

```

H95852
hypothetical fusion protein [imported] - Sinorhizobium meliloti (strain 1021) megaplaasm1
C/Species: Sinorhizobium meliloti
C/Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 31-Dec-2004
A/Accession: H95852
R/Finan, T.M.; Weidner, S.; Wong, K.; Buhrmeester, J.; Chain, P.; Vorholter, F.J.; Hernan
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A/Title: The complete sequence of the 1,683-kb pSymb megaplaasmid from the N2-fixing endo
A/Reference number: A95842; MUID:21396508; PMID:11481431
A/Accession: H95852
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-658 <KIR>
A/Cross-references: UNIPROT:Q92X74; UNIPARC:UPI00000CB3E6; GB:AL591985; PIDN:CAC48488.1;
R/Experimental source: strain 1021, megaplaasmid pSymb
R/Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubier,
pel, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A/Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
heault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
A/Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A/Reference number: A96039; MUID:21368234; PMID:11474104
A/Contents: annotation
C/Genetics:
A/Gene: SMD20088
A/Genome: Plasmid
C/Superfamily: uncharacterized conserved protein with erythromycin esterase domain

```

```

Query Match      25.5%; Score 894.5; DB 2; Length 658;
Best Local Similarity 36.7%; Pred. No. 8.2e-50;
Matches 255; Conservative 81; Mismatches 269; Indels 89; Gaps 23;

Qy      36 RDQGDVVLGLARGGLVAVMEVAAALAPLDAFYVKLGARGHDEFAY--GALASGRV 93
Db      6 RARRMTTSHLSRGIIDRHVEMGVPREAFV---DPGFEEFVDSALSIHQGT 60
Qy      94 VNDVAVGLRTTPOQLR-----DIAREG--RELRRSAY-----RGRPP 133
Db      61 ISQPYIYALMIERAEVPGDTVEIGSGSAAVLSRIAAYVTIERHAGLAVALRRF 120
Qy      134 TDTT-GKTVIVVDGLA--TGASFAV-----QALRDAQPAQIVAVPADEST 180
Db      121 AELRAYGNIDRVAGDGPAPGFDALIVAAAGPEIIPHALKEQDLDGSHLIVPGPPE 180
Qy      181 CREPAGLVDDVVCATMTPTPLA-----VGSFMDRQVTVDEVKRLLTPTGSPSLR 233
Db      181 QRLMK--VTRVNAITTFEEDQDGVGVFVPLVGEYMW-----HEBRAOSRAAPRAPTLPE 232

```

```

Qy      234 PAASTADVLRRVAIDAPGVPTHEVLAEIVGDARIVLIESSHGTHIEFQARAAMTQWL 293
Db      233 LVAE-AAEPJBDL-----DDPAFGLLDFRPAQRRIYLLGEASHTGEFRRAPAAITRRL 285
Qy      294 IEEKFGAVAAEADWPAYVNRVVRGLGEDTNADEALSGFERPPAMMENTVVRDVEV 353
Db      286 IEEHGFALVVEADWPAAVDRVYHRRQGIKLD--AFQRFPTVMNRREVMDEVEM 342
Qy      354 LRTNCRYESGALROAGFYGLDLSLHRSIOEVLSTLDKVDPRAPARAPARACFDHACA 413
Db      343 MREHNGR--KSLDPRAQFYGLDIYNNKGSISALVRIYDEIDPEAAVAARRYGLT--PWQ 399
Qy      414 DCGQAYFAAFAAGPSCEREAVEQLVDVORNALAYARODGLADELFYAQONAAQTVRD 473
Db      400 NEBSTYGRAMTATGFRKCEBAVVRQCCELLEKOLEAGRDGSG---DELLDAVONARLVAS 455
Qy      474 AEVYRAMFSGRVTSMNLRDQMAQTGLSLTLHLDRHLDPAPARIVVAHNSHVGDAARAT 533
Db      456 AERYRIMYYAGADSNMNRDTHMFEITLHL-----NARGARKAVVAHNSHIGDARHT 510
Qy      534 EV-WADGQLTLAGQIVRERYGDSRSIGFSTYGTATYAASEMGIAQRKAVRPALHSGVEE 592
Db      511 DMGARRELINICQLCRERFEDQALIGFTHGHVAAASDMNDMEVKYIRPBLESGYER 570
Qy      593 LPHQT-ADSFVLS---ARLSRDAEAPLDVRLGRAIGVYLPATERQSHYHVRPADQF 647
Db      571 VMHDSGVGCEFLIDPARHARLRDGLKPL-----LERFGIVRDPDTERFHYHAASLPQCF 626
Qy      648 DAMIHIDQTRALEPLVTSRWIAGENPEYPTGL 681
Db      627 DGFVWFQDTPVSEPL--GAQHIRAGVDPTRPGL 658

```

RESULT 3

```

T48790
hypothetical protein J3E11.380 [imported] - Neurospora crassa
C/Species: Neurospora crassa
C/Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 31-Dec-2004
A/Accession: T48790
R/Schulte, U.; Aign, V.; Hohelsel, J.; Brandt, P.; Fattmann, B.; Holland, R.; Nyekatura,
submitted to the Protein Sequence Database, April 2000
A/Reference number: Z24541
A/Accession: T48790
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-554 <SCH>
A/Cross-references: UNIPROT:Q9P6X8; UNIPARC:UPI000017B43F; EMBL:ALJ53620; GSPDB:GN00112;
A/Experimental source: cosmid contig J3E11; strain 74
C/Genetics:
A/Gene: NCSP:J3E11.380
A/Map position: 2
A/Intons: 346/1
C/Superfamily: uncharacterized conserved protein with erythromycin esterase domain

```

```

Query Match      22.2%; Score 780.5; DB 2; Length 554;
Best Local Similarity 38.8%; Pred. No. 1.4e-42;
Matches 192; Conservative 72; Mismatches 174; Indels 57; Gaps 17;

Qy      219 RLLATPTAGP--SLRR-----PAASTADVLRRVAIDAPGVPTHEVLA--ELVGDAR 259
Db      85 ROSTPQPAPSPFLILRRHVHVPFSLSDAD-----DPKQDIFRYFDSFSSAKV 132
Qy      270 VLIAGESHGTHIEFYQARAAMTQWLIEKFGVAEADWPDAVNRVYR---GLGE--- 323
Db      133 LLIIGDASHGISEFYARFALRTYMIHHSINITYACEADMPDASIDRYVRRRPGPRAT 192
Qy      324 -DTNADALSG---FERPPAMMNRVTVVDPEVMLRTNRQRYESGALROAGFYGLDLYS 378
Db      193 ISEBAERARAGRPRPSFMRFRMMWRKETHDFEWLRCKNMKGDM--KKEATGFGYGLDLYS 251
Qy      379 LHSIOEVLSTLDKVDPRAPARAPARACFDHAC--ADCGQAYFAAFAAG---PSCER 433

```


Query Match 4.5%; Score 159; DB 2; Length 446;
 Best Local Similarity 21.1%; Pred. No. 0.0092;
 Matches 92; Conservative 73; Mismatches 174; Indels 96; Gaps 20;

QY 260 LAELVDGARIVLIGSSHGTHFEYQARAATQMLIEKGFAGVAADMDPAVRVRYR 319
 DB 59 LKTVLKGRIYQVIGSTTGAGINATKVMIKYLHEIGYDVLARESGFP----- 108
 QY 320 GLGEBTNDDELSCGERPPAMWRTV-----VRFVEMLRFRNRQRYSG-ALRQAGFYG 373
 DB 109 ----DTNA-SYLNMDQLTPKSTKNSIYAVWHTEDVELFDYMKQEKGDPLITGF-- 161
 QY 374 LDVLSLHRSIQ-EVTSYLDKVDPRRAARARAYACFDHACADDGQAGPAAFGAGPSCE 432
 DB 162 -DIQMKNSFNVAATQWKAADPEKALLSGSENDFTLVDSNTFDEFS----- 210
 QY 433 REAVEQLVDVQRNALAYAR-----QDGLAEDELFYAQNQAQTVR--DAEYYRAMESG 484
 DB 211 -QKKEKLVKNYQKLIKFTKTHASELKENLPKEPKAYEMPMHSLQLRIDMETYMLEMKE 269
 QY 485 RVT-----SNVLBDQHNAGTIGSLTLHDLRLDAPPAIVVMANSHV--GDARAT 533
 DB 270 KLEBYPENIEDFSFMRDRMAAEQFWVADTL---YPKKCIYWGHNHRLRKQNTQMI 324
 QY 534 EVMADGOL---TIGQIVBERYGDSESRISGFSTYGTVAASEMGSLAQKAVRPAL----- 586
 DB 325 KQWV--QLNGPFWGDLPERLKKQYTYTIGIYASGLSDSD-----NKTVPKPTSPPP 376
 QY 587 HGSVEELFHQTADSFVARSRLDAEAP---LDVRLGRAIGV--VYLPAERQSHYLV 641
 DB 377 SGLLEALL-----KADRPVAVFDPLHTKNKKGTSWMTPTPTALYGVMEB 422
 QY 642 RP--ADQPDAMIHD 654
 DB 423 QMILKEQYDGVIMLE 437

RESULT 15

F64671 conserved hypothetical protein HP1214 - Helicobacter pylori (strain 26695)
 C/Species: Helicobacter pylori
 C/Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 09-Jul-2004
 C/Accession: F64671
 R/Tomb, J.F.; White, O.; Kestavange, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khairak, H.G.; Glodek, A.; McKenney, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L. Nature 388, 539-547, 1997
 A/Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Frazer, C. A/Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
 A/Reference number: A64520; MUID:97394467; PMID:9252185
 A/Accession: F64671
 A/Status: preliminary; nucleic acid sequence not shown; translation not shown
 A/Molecule type: DNA
 A/Residues: 1-240 <TOM>
 A/Cross-references: UNIPROT:O25813; UNIPARC:UP100000C0903; GB:AE000627; GB:AE000511; NID C/Genetics:
 A/Start codon: GTG

Query Match 4.5%; Score 157; DB 2; Length 240;
 Best Local Similarity 21.5%; Pred. No. 0.0054;
 Matches 50; Conservative 49; Mismatches 132; Indels 2; Gaps 2;

QY 2 LMTAADVTRSPRRVFRDRERAGVLAELLAAYRDQPDVIVLGLARGLPVAMEVAAL 61
 DB 7 LNTDPSHTIDIEGKRFINEEDALNTLINEIHTRHIDLKDSITMLALSPNALYLAHALAQKF 66
 QY 62 HAPLDAFVYKLGAGCHDEFAVGALASGRVYVNDVYRGLRITPQQLRDIAEREGRE-L 120
 DB 67 GATYDILFLEPLIADLNSKCEIALVSESMDIVMNSLINSFDITLDYVYGEAKRAYEEDI 126
 QY 121 LRRESAYGERPPTDITGKTVIVDDGLATGASMPAAVQALRDAQPAQIVIAVPAPEST 180

DB 127 LSHIYQKGNAIKSLDKNIFIVDRGIETGPRAGLGVQTCLEKCEODIYILTPVAQNV 186
 QY 181 CREPAGLVYDVVCAATMPTPLAVGESFMDPROVDEEYRRLATPTAGPSLR 233
 DB 187 AQLESICDGVISYRPECFVSEHHTYELKRLSNEVEKYLGAANN-PNLKK 238

Search completed: March 23, 2006, 05:26:30
 Job time : 29.0938 secs

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM profile - protein search, using sw model

Run on: March 23, 2006, 04:56:43 ; Search time 174.174 Seconds
(without alignments)

2758.529 Million cell updates/sec

Title: US-10-617-038-23

Perfect score: 3508
Sequence: 1 VLMTAAADVTRSPRRVFRD.....LEVTSRWIAGENPPTPTGL 681

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot_05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3505	99.9	681	053475_MYCTU	053475 mycobacteri
2	3022	86.1	606	077241_MYCBO	077241 mycobacteri
3	1974	56.3	390	050822_MYCTU	050822 mycobacteri
4	1537	43.8	385	050819_MYCTU	050819 mycobacteri
5	1160.5	33.1	495	05NMQ1_AZOSR	05NMQ1 azocarus sp
6	1120.5	31.9	448	04H703_9DEIO	04H703 deinococcus
7	956.5	27.3	679	089197_BRAJA	089197 bradyrhizob
8	919	26.2	653	06W168_RHISM	06W168 rhizobium s
9	894.5	25.5	658	092X74_RHIME	092X74 rhizobium m
10	848	24.2	456	06MCX4_PARUM	06MCX4 paracitlanyd
11	840.5	24.0	455	068HX2_PSEBP	068HX2 pseudomonas
12	816	23.3	445	04IWA2_AZOVI	04IWA2 azotobacter
13	816	23.3	453	05B585_EMENT	05B585 emmentillus
14	811.5	23.1	463	04WJ15_ASPFU	04WJ15 aspergillus
15	800.5	22.8	455	04I713_GIBZE	04I713 gibberella
16	780.5	22.2	474	09P6X8_NEUCR	09P6X8 neurospora
17	780.5	22.2	554	07RUH3_NEUCR	07RUH3 neurospora
18	773.5	22.0	443	04NE69_9MICC	04NE69 arthrobacte
19	710.5	20.3	428	06MM92_BDEBA	06MM92 bdellovibri
20	678	19.3	352	073YR8_MYCBO	073YR8 mycobacteri
21	668	19.0	352	077YV6_MYCBO	077YV6 mycobacteri
22	668	19.0	352	066232_MYCTU	066232 mycobacteri
23	540	15.4	221	08V113_ANASP	08V113 anabaena sp
24	538	15.3	221	08PW22_METMA	08PW22 methanosaic
25	537	15.3	235	063X22_BURPS	063X22 burkholderi
26	538	15.3	220	04H704_9DEIO	04H704 deinococcus
27	527	15.0	230	088HX3_PSEGP	088HX3 pseudomonas
28	520	14.8	241	07N1L5_GIOVI	07N1L5 giobacter
29	483	13.8	215	04IWA5_9BURK	04IWA5 burkholderi
30	479	13.7	218	083E61_COXBU	083E61 coxiella bu
31	471	13.4	216	05WXC9_LEGFL	05WXC9 legionella

32	464	13.2	220	2	05ZWC4_LEGPH	05ZWC4 legionella
33	464	13.2	228	2	04IYW0_AZOVI	04IYW0 azotobacter
34	463	13.2	216	2	05X607_LEGPA	05X607 legionella
35	431	12.3	443	2	053768_MYCTU	053768 mycobacteri
36	431	12.3	443	2	07URJ3_MYCBO	07URJ3 mycobacteri
37	431	12.3	451	2	07DPW2_MYCTU	07DPW2 mycobacteri
38	423	12.1	441	2	06MGV6_BDEBA	06MGV6 bdellovibri
39	418.5	11.9	206	2	05YQX9_NOCFA	05YQX9 nocardia fa
40	415.5	11.8	258	2	050111_MYCUE	050111 mycobacteri
41	415	11.8	510	2	04NB36_9MICC	04NB36 arthrobacte
42	414.5	11.8	248	2	04IWA3_AZOVI	04IWA3 azotobacter
43	413.5	11.6	235	2	0989T9_RHIZO	0989T9 rhizobium l
44	406	11.6	236	2	092NP2_RHIME	092NP2 rhizobium m
45	404	11.5	79	2	077242_MYCBO	077242 mycobacteri

ALIGNMENTS

RESULT 1
ID 053475 MYCTU PRELIMINARY; PRT; 681 AA.
AC 053475_07D7L3;
DT 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DT 13-SEP-2005 (TREMBlrel. 31, Last annotation update)
DE Hypothetical protein.
GN OrderedLocustNames=MT2089, Rv2030c;
OC Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium tuberculosis complex.
OX NCBI_TaxID=1773;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230; DOI=10.1038/31159;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C.M.,
RA Harris D.E., Gordon S.V., Eigmeier K., Gae S., Barry C.E. III,
RA Tekala F., Badcock K., Basham D., Brown D., Chillingworth T.,
RA Connor R., Davies R.M., Devlin K., Felwell T., Gentles S., Hamlin N.,
RA Holtroyd S., Hornby T., Jagers K., Krogh A., McLean J., Moule S.,
RA Murphy L.D., Oliver S., Osborne J., Quail M.A., Rajandream M.A.,
RA Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Disciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."
RL Nature 393:537-544(1998).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CDC 1551 / Oshkosh;
RX MEDLINE=22206494; PubMed=12218036;
RX DOI=10.1128/JB.184.19.5479-5490.2002;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J.D., DeBoy R.T., Dodson R.J., Gwyn M.L., Haft D.H.,
RA Hickey B.K., Kolony J.F., Nelson W.C., Umayan L.A., Ermolaeva M.D.,
RA Salzberg S.L., Delcher A.L., Uterback T.R., Weidman J.F., Khouri H.M.,
RA Gill J., Mikha A., Bishai W., Jacobs W.R. Jr., Venter J.C.,
RA Fraser C.M.;
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains."
RT J. Bacteriol. 184:5479-5490(2002).
CC -1- SUBUNIT: Homodimer (By similarity).
CC -1- SUBUNIT: Homodimer (By similarity).
CC -1- SUBUNIT: Homodimer (By similarity).
EMBL: BX842578; CAA17244.1; -; Genomic DNA.
EMBL: AB000516; AAK46368.1; -; Genomic DNA.
DR PIR; E70942; E70942.
DR TIGR; MT2089; -.
DR Tuberculist; Rv2030c; -.
DR GO; GO:0016789; F:carboxylic ester hydrolase activity; IEA.
DR GO; GO:0016757; F:transferase activity, transferring glycosyl. . .; IEA.
DR GO; GO:0009116; P:nucleoside metabolism; IEA.
DR GO; GO:0006166; P:purine ribonucleoside salvage; IEA.

Query Match	99.9%;	Score 3505;	DB 2;	Length 681;
Best Local Similarity	99.9%;	Pred. No. 36-206;		
Matches 680;	Conservative 1;	Mismatches 0;	Indels 0;	Gaps 0;
DR GO:0046677; P:response to antibiotic; IDA.				
DR InterPro; IPR007815; Erythroesteras.				
DR InterPro; IPR000836; PRTtransferase.				
DR Pfam; PF05139; Erythro_esteras; 1.				
DR Pfam; PF00156; Pribosyltran; 1.				
KW Complete proteome; Glycosyltransferase; Hypothetical protein;				
KW Transferase.				
SEQUENCE 681 AA: 74930 MW: 82046ACCS550947 CRC64;				
QY 1 VLMTAADVTRRSPPRRVPRDRREAGRVIAELIAAYRDPDVTGLGAGLPVAMEVAA 60				
DB 1 MLMTAAADVTRSPRRVPRDRREAGRVIAELIAAYRDPDVTGLGAGLPVAMEVAA 60				
QY 61 LHAPLDAVVRVLGAGHGFDEFAVGALAGGRVYVNDVVRGRLTPQQLRIAREGREL 120				
DB 61 LHAPLDAVVRVLGAGHGFDEFAVGALAGGRVYVNDVVRGRLTPQQLRIAREGREL 120				
QY 121 LRRESAYGERPPTITGKTVLVVDGLATGASMPAAVQALRDAPQIVIAVPAPEST 180				
DB 121 LRRESAYGERPPTITGKTVLVVDGLATGASMPAAVQALRDAPQIVIAVPAPEST 180				
QY 181 CREPAGLVDDVCACTPPTPLAVGESFMDPRQVTEDEVRLLATPTAGPSLRPAASTAA 240				
DB 181 CREPAGLVDDVCACTPPTPLAVGESFMDPRQVTEDEVRLLATPTAGPSLRPAASTAA 240				
QY 241 DVLRRVALDAGGAVPTHEVLAEVGDARIVLIGESSHGTHERYQARAAMTWLIEKGF 300				
DB 241 DVLRRVALDAGGAVPTHEVLAEVGDARIVLIGESSHGTHERYQARAAMTWLIEKGF 300				
QY 301 AVAAEADPDPAYRVVRYRGCEPTNADAEALSGFERPPAMMMRNTVVADEPVEMLRTRNQR 360				
DB 301 AVAAEADPDPAYRVVRYRGCEPTNADAEALSGFERPPAMMMRNTVVADEPVEMLRTRNQR 360				
QY 361 YESGALROAGFYGLDLYSLHSISQEVISYLDKVPBRAAARARAYACPDHACADGQAYG 420				
DB 361 YESGALROAGFYGLDLYSLHSISQEVISYLDKVPBRAAARARAYACPDHACADGQAYG 420				
QY 421 FFAAFGAGPSCEREAVEQLVQQRNALAAYARODGLADELFYAQQNAQTVRDASEVYRA 480				
DB 421 FFAAFGAGPSCEREAVEQLVQQRNALAAYARODGLADELFYAQQNAQTVRDASEVYRA 480				
QY 481 MFSGAVTSMNLRDQMAQTLSLTLHLDRLHLDAPARLVVMAHNSHVGDARATEVMADGQ 540				
DB 481 MFSGAVTSMNLRDQMAQTLSLTLHLDRLHLDAPARLVVMAHNSHVGDARATEVMADGQ 540				
QY 541 LTLGQIVNERVGDSESRISGFSYTGTVTAASEWGSLAQRKAVRALHGSVEELFHOQTDS 600				
DB 541 LTLGQIVNERVGDSESRISGFSYTGTVTAASEWGSLAQRKAVRALHGSVEELFHOQTDS 600				
QY 601 FLVSAKLSRDAEAPLDVRLGRLGAVGVVLPATEROSHYLAVRPADQPDAMIHIQTRALE 660				
DB 601 FLVSAKLSRDAEAPLDVRLGRLGAVGVVLPATEROSHYLAVRPADQPDAMIHIQTRALE 660				
QY 661 PLEVTSRWIAGENPETPTGL 681				
DB 661 PLEVTSRWIAGENPETPTGL 681				
RESULT 2				
Q7TZ41_MYCBO PRELIMINARY; PRT; 606 AA.				
AC Q7TZ41; ID Q7TZ41; MYCBO PRELIMINARY; PRT; 606 AA.				
DT 01-OCT-2003 (TREMBLrel. 25, Created)				
DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)				
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)				
DN Hypothetical protein Mb2056c.				
SN OrderedLocustNames=Mb2056c;				
OS Mycobacterium bovis.				
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;				

CC Corynebacteriineae; Mycobacteriaceae; Mycobacterium;
 OC Mycobacterium tuberculosis complex.
 OX NCBI_TaxId=1765;
 RP [1]
 RN NUCLEOTIDE SEQUENCE.
 RC STRAIN-AF2122/97;
 RK MEDLINE=22709107; PubMed=12788972; DOI=10.1073/pnas.1130426100;
 RA Garnier T., Eglmeier K., Camus J.-C., Medina N., Mansoor H.,
 RA Harris M., Duboy S., Grondin S., Lacroix C., Monsempé C., Simon S.,
 RA Harris B., Actin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
 RA Parthill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
 RT "The complete genome sequence of Mycobacterium bovis";
 RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
 CC -1- SUBUNIT: Homodimer (By similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 DR EMBL, BX248341; CAD96909.1; -; Genomic DNA.
 DR GO: GO:0016789; F:carboxylic ester hydrolase activity; IEA.
 DR GO: GO:0016757; F:transferase activity, transferring glycosyl. .; IEA.
 DR GO: GO:0009116; P:nucleoside metabolism; IEA.
 DR GO: GO:0006166; P:purine ribonucleoside salvage; IEA.
 DR GO: GO:0046677; P:response to antibiotic; IEA.
 DR InterPro: IPR007815; Erythro_esterae.
 DR InterPro: IPR000836; PRTasefaae.
 DR Pfam: PF05139; Erythro_esterae; 1.
 DR Pfam: PF00156; Pribosyltran; 1.
 KM Complete proteome; Glycosyltransferase; Hypothetical protein;
 KW Transase.
 SQ SEQUENCE 606 AA; 66540 MW; C9E2278908BF4747 CRC64;

Query Match	Similarity	Score	DB 2,	Length	606;
Best Local	56.1%;	99.8%;	Fred. No. 1e-176;		
Matches	586;	Conservative	1;	Indels	0;
				Gaps	0;
QY	1	VLAMTAAADVTRSSRRVPRDRREAGRYLAETLLAAYRDPQDVIVTGLARGLPVAWEVAA	60		
DB	1	MLMTAAADVTRSSRRVPRDRREAGRYLAETLLAAYRDPQDVIVTGLARGLPVAWEVAA	60		
QY	61	LHAPLDAFVVRKLGAPGHDEFVAGALASGGHVVVNDVDVVRGLRTTPOOLRIAREGSEL	120		
DB	61	LHAPLDAFVVRKLGAPGHDEFVAGALASGGHVVVNDVDVVRGLRTTPOOLRIAREGSEL	120		
QY	121	LRRSARGEPRPDITGKTVIVVDDGLATASMPAAVQALRDQOPQIVIAVPAAPST	180		
DB	121	LRRSARGEPRPDITGKTVIVVDDGLATASMPAAVQALRDQOPQIVIAVPAAPST	180		
QY	181	CRFEGAGLVDDVVCATMPTPFLAVGESFMDPRQVDEEVRLLATPTGSPSLRPAASTAA	240		
DB	181	CRFEGAGLVDDVVCATMPTPFLAVGESFMDPRQVDEEVRLLATPTGSPSLRPAASTAA	240		
QY	241	DVLRVVAIDAFGSPVTHEVTLAELVGDARIIVLIGSSHCTHEFYQARAAMTOMLIEKCFG	300		
DB	241	DVLRVVAIDAFGSPVTHEVTLAELVGDARIIVLIGSSHCTHEFYQARAAMTOMLIEKCFG	300		
QY	301	AVAAEADMPDXYRNVRYRGLGEPTNDAEALSGERPAMMMRTTVRDVFEMLRTRNOR	360		
DB	301	AVAAEADMPDXYRNVRYRGLGEPTNDAEALSGERPAMMMRTTVRDVFEMLRTRNOR	360		
QY	361	YESGALRQAGFYGLDLYSLHRSIOEVI SYLDKVDPRPAAARARAFYACFDHACADGQAYG	420		
DB	361	YESGALRQAGFYGLDLYSLHRSIOEVI SYLDKVDPRPAAARARAFYACFDHACADGQAYG	420		
QY	421	FAAAFAGSPGSEBRANVEQLVDVQNNALAYARQDGLADELFYAQONNAQYTRDAEYVYRA	480		
DB	421	FAAAFAGSPGSEBRANVEQLVDVQNNALAYARQDGLADELFYAQONNAQYTRDAEYVYRA	480		
QY	481	MFSGRVYSMNLRDHMAQTGLSLTHLDRHLDAEPARIIVWAHNSHVGDAEATEVMADGQ	540		
DB	481	MFSGRVYSMNLRDHMAQTGLSLTHLDRHLDAEPARIIVWAHNSHVGDAEATEVMADGQ	540		
QY	541	LTTGQIVRERYGDESRSIGFSTYTGTVTAASEWGIAQORAVRPAH	587		
DB	541	LTTGQIVRERYGDESRSIGFSTYTGTVTAASEWGIAQORAVRPAH	587		

Query Match	56.3%	Score 1974;	DB 2;	Length 390;
Best Local Similarity	97.7%	Pred. No. 9.2e-113;		
Matches 376; Conservative	2;	Mismatches 7;	Indels 0;	Gaps 0

Qy	65	RALEPLEVTSRWIAGENPETYPPTGL	681
Db	366	RALEPLEVTSRWIAGENPETYPPTGL	390

OC	<i>Mycobacterium tuberculosis</i> complex.
OX	NCBI_TaxID=1773;
RN	[1]

Query Match	43.8%	Score 1537;	DB 2;	Length 385;
Best Local Similarity	99.0%	Pred. No. 5.3e-86;		
Matches 306; Conservative	1;	Mismatches 2;	Indels 0;	Gaps 0;

QY	301	AVAAEADWP	309
Db	301	AVAAEATGP	309

DR Pfam; PF05139; Erythro_esterase; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 495 AA; 55705 MW; D8D718A9BFF77A CRC64;

```
Query Match      33.1%; Score 1160.5; DB 2; Length 495;
Best Local Similarity 50.7%; Pred. No. 8.3e-63;
Matches 243; Conservative 64; Mismatches 135; Indels 37; Gaps 10;

QY 229 PSLEPRPAASTA-----ADVLRRVAI---DAPGVPTHEVLAELVDARIVL 271
DB 28 PGRRTPRRSRSPGKKEKTKMERDITRAVVASAIEPLKDGTRMDLALLIGDARIVL 87
QY 272 IGESSHGTHFEFYQAAMTQWLIEKFGVAEAADWDAYVNRVYRGLEEDTNADEAL 331
DB 88 LGEASHGTHFEFYAARAGITQRLITEKFDVAVEADWDPSLRASRYAGSGDDADANAL 147
QY 332 SGFERFPAMWRNTVVDVFEWMLTRNQRYSGLRQAGFYGLDLSLHNSIQEYISYLD 391
DB 148 GGFQRFPRMMWRNTEVAFAFIEMLRQVNS--SRGADERVGFGLDLSYSLRASMDEVVRYLE 205
QY 392 KVDPRAARARARACPDHACADQOAYGFAAAGAGSGCERAVEOLVDVQRNALAYAR 451
DB 206 TVDEBAARARARACPDH-IAEDPQRIGVATTTGVLVHCHEREVLRQLAELTGDPERYLH 264
QY 452 QDGLAEBELFYAQQNQTVADEAVYRYAMFSGRYTSNNLRDQMAQTLGSLTLHLDRHL 511
DB 265 HDGIAAADDELFYAQONARVANAEEAYRYSMFAGHESNNRDSMAETLEMLLGHLDSS- 323
QY 512 DAPARIVYMAHNSHVGDAATEVWADQQLTGQIVNERGYDE-SRSIGESTYTYTTAA 570
DB 324 RGRPAKIAVMAHNSHLDGARTEWGEAGEHNLGQIVRQFGEAELARLGFTHAGTYAA 383
QY 571 SEMGIIQRKAVRPAALGVSVELFHQT-ADSFVLSA-----RLSDAEAPLVDVNLGRAIG 625
DB 384 SDWQSPFLKVRKVRSHSPESFERVHDIGVERFLPLASDRVPLRKRER-----RLERRAIG 436
QY 626 VVYLPTATERSHYLHVPRADQFDAMIHIDQTRALEPLEVTSRMTAGENP---ETVPTGL 681
DB 437 VYLPGERVSHYFYADIAKQFDVAVIHDRTRALTPLDPSITMWHGSSPPEQMERTYPSGM 495

RESULT 6
Q4HTU3_9DEIO PRELIMINARY; PRT; 448 AA.
AC 0447U3;
DT 13-SEP-2005 (TReMBLrel. 31, Created)
DT 13-SEP-2005 (TReMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TReMBLrel. 31, Last annotation update)
DE Erythromycin esterase.
GN ORFNames=dgeODRAFT_0443;
OS Deinococcus geothermalis DSM 11300.
OC Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;
OC Deinococcaceae; Deinococcus.
OX NCBI_TaxID=319795;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=DSM 11300;
RG US DOE Joint Genome Institute (JGI-ORNL);
RA Larimer F., Land M.;
RT "Annotation of the draft genome assembly of Deinococcus geothermalis
RT DSM 11300."
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL, AAHE01000008; EMBL2343.1; -, Genomic DNA.
SQ SEQUENCE 448 AA; 5095 MW; 6E3EEFE136D64750 CRC64;
```

```
Query Match      31.9%; Score 1120.5; DB 2; Length 448;
Best Local Similarity 53.1%; Pred. No. 2.1e-60;
Matches 234; Conservative 62; Mismatches 134; Indels 11; Gaps 8;

QY 243 LRRVADPAGVPTHEVLAELVDGARIVLIGESSHGTHFEFYQAAMTQWLIEKFGAV 302
DB 15 LRVDARPLTGAADYDLDLDCIGNARFVLIGESSHGTHFEYRRARLTMILIEKFTAV 74
QY 303 AAEADWDAYVNRVYRGLEEDTNADEALSGERFPAMWRNTVVDVFEWMLTRNQRYE 362
DB 75 AVEADWDAYVNRVYRGQADKSALAEALDFQRFPRMMWRNBDVOAFVSSLRDHNR-S 133
QY 363 SGALRQAGFYGLDLSLHNSIQEYISYLDKVDPRRAARARACPDHACADQOAYGFA 422
DB 134 PGA--PVGFGLDLSLHNSMAVVELETVDDEAARARQGCDFQ--FGENQOAGYA 190
QY 423 AAGAGSGCERAVEOLVDVQRNALAYARDDGLAEDELFYAQQNQTVADEAVYRYAMF 482
DB 191 TESGRREPCDDAAVQQLLELQRRA--QESGFLAEDERFYAQONARLAKNAETTYRAMF 248
QY 483 SGRVTSNNLRDQMAQTLGSLTLHLDRHLDAPPARIVYMAHNSHVGDAATEV-WADQQL 541
DB 249 RGRDESWSLDAMWAETLEALVEGSR--QGRPOKIVYMAHNSHLDGARASEMGWLRGEL 306
QY 542 TLQOIVNERYGDSRSIGESTYTYTTAASEKGIQRKAVRPAALGVSVELFHQTRAD-S 600
DB 307 NVQQLAERPRGFTFYIGQSTHGTGYAADDWDEPARVARPALPQSVVDLLHVEGEAA 366
QY 601 FIVSARLSRDAEAPLVDVNRGIRAGVYVLPATERSHYLHVPRADQFDAMIHIDQTRALE 660
DB 367 YMLDKRENPATYGLRKRERQRFQIYIRPETERMSHYVTRLSMDVALLPFDERSAVV 426
QY 661 PLEVTS-RWLAGENPETYPTG 680
DB 427 PLDATAGTEPEGEVPTDFTPTG 447

RESULT 7
Q089L97_BRAVA PRELIMINARY; PRT; 679 AA.
AC 089L97;
DT 01-JUN-2003 (TReMBLrel. 24, Created)
DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Protein-L-isoaspartate O-methyltransferase.
GN OrderedLocuNames=b114651;
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
OX NCBI_TaxID=375;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=USDA 110;
RG MEDLINE=22484998; PubMed=12592725;
RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,
RA Sasamoto S., Watanabe A., Idegawa K., Iriyuchi M., Kawashima K.,
RA Kohara M., Matsumoto M., Shimo S., Tsurunaka H., Wada T., Yamada M.,
RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RT Bradyrhizobium japonicum USDA110."
RL DNA Res. 9:189-197(2002).
CC -1- FUNCTION: Catalyzes the methyl esterification of L-isoaspartyl
CC residues in peptides and proteins that result from spontaneous
CC decomposition of normal L-aspartyl and L-asparaginyl residues. It
CC plays a role in the repair and/or degradation of damaged proteins
CC (by similarity).
CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + protein L-
CC isoaspartate = S-adenosyl-L-homocysteine + protein L-isoaspartate
CC alpha-methyl ester.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (by similarity).
DR EMBL, BA000040; BAC49916.1; -, Genomic DNA.
DR HSP; Q81ZK3; 1UG1.
```

DR GO:0016789; F:carboxylic ester hydrolase activity; IEA.
 DR GO:0004719; F:protein-L-isoaspartate (D-aspartate) O-meth. .; IEA.
 DR GO:0006464; P:protein modification; IEA.
 DR GO:0046677; P:response to antibiotic; IEA.
 DR InterPro: IPR007815; Erythro_esterases.
 DR InterPro: IPR00682; PCMT.
 DR InterPro: IPR00051; SAM_bd.
 DR Pfam: PF05139; Erythro_esterases; 1.
 DR Pfam: PF01135; PCMT; 1.
 DR TIGRFAMs: TIGR00080; pmtc; 1.
 DR PROSITE: PS01279; PCMT; 1.
 KM Complete proteome; Methyltransferase; Transferase.
 SQ SEQUENCE 679 AA; 75244 MW; AFE5B82707429338 CRC64;

Query Match 27.3%; Score 956.5; DB 2; Length 679;
 Best Local Similarity 37.3%; Pred. No. 3.8e-50;
 Matches 259; Conservative 72; Mismatches 207; Indels 157; Gaps 22;

QY 24 AGRVLAELAAVRDQDVIVIGLARGGLPVAMEVAAL-----H 62
 DB 105 AAATVSEIAA---DYVTVRLG-----PLAEAAATLTGLGYDVAHVLAGDGTGMDEH 155
 QY 63 APLDAFVRKLGAPGHDPEFVAGALASGGRVV-----NDVVRGLRITPQQLR--DIA 113
 DB 156 APYDAIVV-AAGGPVPALKEQLKIGRLVIVPGADQRTQELVAVVRLSKDEFSESDIA 214
 QY 114 EREGSEILRR--SAYRGERPPTDTGKTVIVDDGLATGASMPAAVQALDAQPAQVI 171
 DB 215 DVRFVPLGEGGMAAKETAP-----ARAAAR----- 242
 QY 172 AVPAPESTCEBFAGLVDDVVCATMPPLAVGSEFMDFROVTDDE--VRRLATPTAGP 229
 DB 243 --PVAP-----DETLVRINADAESEFP 263
 QY 230 SLRRPASTADYLRVAIDAAGVPTHEVLAELVGDARIVLIGSSHGTEFYQARAA 289
 DB 264 SIE-----AADL-----SPLMERIGSARVVLGATIGTSEFPYMRERI 302
 QY 290 TQWLIIEKGFGVAEAADMPDAYRVNRYVRLGSEDTNADALSGFERFPAMMMRTVVRD 349
 DB 303 TRDLIVKKGFRFVAIEADMPDAARDVHVRH---QYRPESEMTAFARFTMMWMTTEVRD 359
 QY 350 FVEWELRTNQRESGALRQAGFYGLDLSLHRSIQEIVISYDKVDPRAADPARARYACPD 409
 DB 360 FVSWLRKNGTVEN--TRVAHFHGIDLSLYDSIRSVINLYDEVDPASARVARERYGCLT 417
 QY 410 HACADDGAYGFAAFAFGSPCERAVRQVLDVQGNALAAVRODGLAEDLFTAQONAQ 469
 DB 418 -PWQRDPATYGHAAITGSPPTCESDVAAHALTDLAKRRAVVAEHDD---ERFLDAEQAR 472
 QY 470 TVRDAEVYRAMFSGRTVSMILRDOHMAQTGLTTHLRHLADAPARIVVAHNSHYGD 529
 DB 473 LVAAHNERYYRLMYIGSRASWNLRSHPDILKNILA-----FHGSDSAVVAHNSHYGN 527
 QY 530 ARATEWADGQLTLGQIVREERYGDSRSIGFSTYGTVTAASEWGIQAKRAVRPALGSG 589
 DB 528 AAATEWVARGENHIGQHCKAFGAQAYLVGCTHSGTVAASDMGSPMEVTVTRSLPNS 587
 QY 590 VEEFLHQTA--DSFLVASLSDAEP--LDVVRGLRAIGVYVLTPTREQSHYLAHRPADQ 646
 DB 588 YERLCHAHTGLPRFMGLRGPDLCGPEBGLGERLERARAGVYRPELASHYFQASLPDQ 647
 QY 647 FDAMIHIDQTRALBPLEVTSRWIAGENPEYPTGL 681
 DB 648 FDEYIWFDDTRAVTLETA--IAG-LPDMTFPGV 679

RESULT 8
 O6M168 RHISN
 ID O6M168 RHISN PRELIMINARY; PRT: 653 AA.
 AC O6M168_7
 DT 05-JUL-2004 (Tremblrel. 27, Created)
 DT 05-JUL-2004 (Tremblrel. 27, Last sequence update)

DT 05-JUL-2004 (Tremblrel. 27, Last annotation update)
 DE Protein-L-isoaspartate O-methyltransferase (EC 2.1.1.77).
 GN ORNames=NGR00375;
 OS Rhizobium sp. (strain NGR234).
 OG Plasmid megaplasmid 2.
 CC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 CC Rhizobiaceae; Rhizobium.
 OX NCBI_TaxID=394;
 RN [1]
 RP NCULEOTIDE SEQUENCE.
 RC STRAIN=NGR234;
 RA Broughton W.J., Perret X., Staehelin C., Schmitz R.A., Raesch C.,
 RA Liesegang H., Gottschalk G., Streit W.R.;
 RT "Comparative DNA analysis of two large contigs of the Rhizobium sp.
 NGR234 megaplasmid 2.";
 RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
 CC -I FUNCTION: Catalyzes the methyl esterification of L-isoaspartyl
 residues in peptides and proteins that result from spontaneous
 decomposition of normal L-asparaginyl and L-asparaginyl residues. It
 plays a role in the repair and/or degradation of damaged proteins
 (by similarity).
 CC -I CATALYTIC ACTIVITY: S-adenosyl-L-methionine + protein L-
 isoaspartate = S-adenosyl-L-homocysteine + protein L-isoaspartate
 CC alpha-methyl ester.
 CC -I SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC EMBL: AY167479; AA087500.1; -; Genomic DNA.
 DR GO:0016789; F:carboxylic ester hydrolase activity; IEA.
 DR GO:0004719; F:protein-L-isoaspartate (D-aspartate) O-meth. .; IEA.
 DR GO:00000179; F:RNA (adenine-N6,N6)-dimethyltransferase a. .; IEA.
 DR GO:0008649; F:RNA methyltransferase activity; IEA.
 DR GO:0016740; F:transferase activity; IEA.
 DR GO:0006464; P:protein modification; IEA.
 DR GO:0046677; P:response to antibiotic; IEA.
 DR GO:0000154; P:RNA modification; IEA.
 DR InterPro: IPR007815; Erythro_esterases.
 DR InterPro: IPR000682; PCMT.
 DR InterPro: IPR001737; RNA_meth_trans.
 DR InterPro: IPR00051; SAM_bd.
 DR Pfam: PF05139; Erythro_esterases; 1.
 DR Pfam: PF01135; PCMT; 1.
 DR SMART: SM00650; RADc; 1.
 DR TIGRFAMs: TIGR00080; pmtc; 1.
 DR PROSITE: PS01279; PCMT; 1.
 KM Methyltransferase; Plasmid; Transferase.
 SQ SEQUENCE 653 AA; 72564 MW; E2BFF829C764E511 CRC64;

Query Match 26.2%; Score 919; DB 2; Length 653;
 Best Local Similarity 37.4%; Pred. No. 7.3e-48;
 Matches 257; Conservative 66; Mismatches 228; Indels 136; Gaps 23;

QY 6 AADYTRSPRRVF---RDREAGRVLAELAAVYRDQDVIVIGLARGGLPVAMEVAALH 62
 DB 92 AAATVSEIAA---DYVTVRLG-----PLAEAAATLTGLGYDVAHVLAGDGTGMDEH 155
 QY 63 APLDAFVRKLGAPGHDPEFVAGALASGGRVVNDVVRGLRITPQQLRDAEREGRELL- 121
 DB 143 APFADIVV-----AAGGPV-----PRLKEQLTIGGRLLIP 174
 QY 122 --RESAYRGERPPTDTGKTVIVDDGLATGASMPAAVQALDAQPAQVIANPAPES 179
 DB 175 VGRGEQQRLLR---TT-----RTAANRY-----EEEDLGGVLFVP----- 207
 QY 180 TCRFAGLVDDVVCATMPPLAVGSEFMDFROVTVDEVRRLATPTAGSLRRPASTA 239
 DB 208 -----LIGEGNSSESPDHARRV-----PDL-----VAEA 235
 QY 240 ADVLRVAIDAPGCVPTHEVLAELVGDARIVLIGSSHGTEFYQADAAATQWLIIEKGF 299
 DB 236 ADDLPLVT--DGPADPFERFAR---HRIVLGAEATHGTSEFYQADAAVITQRLIERHGF 289
 QY 300 GAVAAEADMPDAYRVNRYVRLGSEDTNADALG---FERFPAMMMRTVVRQVLEWELRT 356
 DB 290 AIVAVEADMPDAAVNVAVVR-----TREAAAGSEPPFORFPAMMMNNAEFAVSWLKT 342

QY 357 RNQRYEGALRQAGFYGLDLYSLHRSIQEIVSYLDKVDPRRAARARAYACFDHACDDG 416
 Db 343 HNG--TGSTORQAGFYGLDLYNMRGSIAYVLSYLDSDRAATVARERYGCLTPEWQEPG 400
 QY 417 QAYGPAAPAGGPGSCERAEVQQLVDQKNAALAYARQDGLADELFTQAQNAQVRAAEV 476
 Db 401 -TYRAVALTSGYRCCEADVIRQCCELLRALDPA-----DQDVAFFDAQVARIATAEK 453
 QY 477 YRRAMFSGRVTSKMLRDQMAQTLGSLTLTLDRHLDAAPARIYVMAHNSHVGDARATEW 536
 Db 454 YRRIMYGGADAMLRDRHMFETLDHLERAD-----PTSKAIYMAHNSHIGDARHTDMG 508
 QY 537 -ADQQLTGVQIVRERYGDESRISGFSTYTGVTVAASEWGIQAKKAVPALHGSVEELFH 595
 Db 509 KLRGELNMGQLCREKYGDVAVLGFGRGTGVAASAWDQMEIKAVRPSLEETVERCCH 568
 QY 596 QTA-DSFLVSKRLSRDAEAPLDVRLGKAIGVYLPAEROSHYLHVRPADQFDMHID 654
 Db 569 LSGERFLDPRKPDVAQSSLSERLERFVIGVIRPESELLSHYARAEISKQFPAAYWFD 628
 QY 655 QTRALEPLEVTSRWIAGENPEYPTGL 681
 Db 629 ETRAVSELPVTAAG--SGHVDPTEPFG 653

RESULT 9

Q92X74_RHIME PRELIMINARY; PRT; 658 AA.
 AC Q92X74_1
 DT 01-DEC-2001 (Tremblrel. 19, last sequence update)
 DT 01-MAR-2004 (Tremblrel. 26, last annotation update)
 DE Protein-L-isoaspartate O-methyltransferase.
 GN OrderedlocusNames=RB0088; ORFNames=SWB20088;
 OS Rhizobium meliloti (Sinorhizobium meliloti).
 OC Plasmid pSym.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Rhizobiaceae; Sinorhizobium.
 NCBI_TaxID=382;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=1021;
 RX MEDLINE=21396508; PubMed=11481431; DOI=10.1073/pnas.161394698;
 RA Finan T.M., Weidner S., Wong K., Buhmeester J., Chain P.,
 RA Vorholster F.J., Hernandez-Lucas I., Becker A., Cowie A., Gouzy J.,
 RA Golding B., Puelher A.;
 RT "The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-
 RT fixing endosymbiont Sinorhizobium meliloti.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9889-9894(2001).
 CC -1- FUNCTION: Catalyzes the methyl esterification of L-isoaspartyl
 CC residues in peptides and proteins that result from spontaneous
 CC decomposition of normal L-aspartyl and L-asparaginyl residues. It
 CC plays a role in the repair and/or degradation of damaged proteins
 CC (by similarity).
 CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + protein L-
 CC isoaspartate = S-adenosyl-L-homocysteine + protein L-isoaspartate
 CC alpha-methyl ester.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (by similarity).
 DR EMBL, AL591985; CAC48488.1; -, Genomic_DNA.
 DR F01; H95852; H95852.
 DR HSBP; Q872R3; 10G1.
 DR GO; GO:0016789; F:carboxylic ester hydrolase activity; IEA.
 DR GO; GO:0004719; F:protein-L-isoaspartate (D-aspartate) O-meth. ., IEA.
 DR GO; GO:0006464; P:protein modification; IEA.
 DR GO; GO:0046677; P:response to antibiotic; IEA.
 DR InterPro; IPR007815; Erythro_esterases.
 DR InterPro; IPR000682; PCMT.
 DR Pfam; PF05139; Erythro_esterases; SAM bd.
 DR Pfam; PF01135; PCMT; 1.
 DR TIGRfam; TIGR00080; pimt; 1.
 DR PROSITE; PS01279; PCMT; 1.

KM Complete proteome, Hypothetical protein; Methyltransferase; Plasmid;
 KW transferase.
 SQ SEQUENCE 658 AA; 73849 MW; C6C58A45F3331CD2 CRC64;

Query Match 25.54; Score 894.5; DB 2; Length 658;
 Best Local Similarity 36.74; Pred. No.2.3e-46;
 Matches 255; Conservative 81; Mismatches 269; Indels 89; Gaps 23;

QY 36 RDQPDVIVLIGLARGSLPVAVEVAALAPLDAFVVRKLGAPGHDEPAV--GALASGRVV 93
 Db 6 RARERMITSHLSRGRIDRRVVEAMGVPREAVV-----DQGEFPAVEDSALSIHQGT 60
 QY 94 VNDVVRGLRITVOQLR-----DIABRG--RELRRESAY-----RGRPP 133
 Db 61 ISQPYIYALMIEBAEVOGPDVTVEIGTSGSYAAVLSRIAAVYTTERRHGLAEVARRP 120
 QY 134 TDTT-GKTVVVDGLA--TGASMPAAV-----QALRDAQVIVAPPAPEST 180
 Db 121 AELRYGNIDVRVGGTRGMPFAGPFDAIYVAGGPEIPHALKEQLDGLHVLVPGPPEE 180
 QY 181 CREFFAGLVDDVCAWTPTPLA-----VGSEFMDFRQVTDDEVRLATTPAGPSLR 233
 Db 181 QRLMK--VTRVNAITTFEEDLGGVRFVPLVGEYGV-----HEBRASRRRPAPTLPE 232
 QY 234 PAASTADVLRRVALIDAPGVPPTHREVLAEVLGDARLYLIGSESHGTHEFYQARAAQTWL 293
 Db 233 LVAE-AAEPLPDL-----DDPAFGLLPDRFAGRRIVLLGEASHGTSEFRARAAITRL 285
 QY 294 IEEKFGVAABADWPDYAVVRNRYVRGLGEDTNADEALSGFERFPAMMNTVVRDFVEW 353
 Db 286 IEEHGFALVAEADWPDAAVDRYVRRRQGITLD--APQGFPTMMNRKREVMDFEY 342
 QY 354 LRTNRQYESGALRQAGFYGLDLYSLHRSIQEIVSYLDKVDPRRAARARAYACFDHAC 413
 Db 343 MREHNRR--RSLPDRAGFYGLDLYNMRGSIAYVLSYLDSDRAATVARERYGCLT-PMQ 399
 QY 414 DDQAYGPAAPAGGPGSCERAEVQQLVDQKNAALAYARQDGLADELFTQAQNAQVRAE 473
 Db 400 NEPSITYGRAAMTGAFFKCEBAVVRQCELEKLEGRQSG--DELDVAVGNALIVAS 455
 QY 474 AEVYRAMFSGRVTSNNLRDOHMAQTLGSLTLTLDRHLDAAPARIYVMAHNSHVGDARAT 533
 Db 456 AERYIIMYTAGADSNMRDTHMFTLHLL-----NAGARSKAVVMAHNSHIGDARHT 510
 QY 534 EV-WADQQLTGVQIVRERYGDESRISGFSTYTGVTVAASEWGIQAKKAVPALHGSVEE 592
 Db 511 DMGAREELNIGQLCERFEDQALIGFGTHGVAAASDMQDMVKPRLPESYSYER 570
 QY 593 LRFQGT-ADSLVS-----ARLSRDAEAPLDVRLGKAIGVYLPAEROSHYLHVRPADQ 647
 Db 571 VMHDSGVGCFLIDPARHARLDGLKPL-----LERFIVYREDTERFSHYAASLPQF 626
 QY 648 DAMIHIDQTRALEPLEVTSRWIAGENPEYPTGL 681
 Db 627 DGFVWPDQTTVPBPL--GAOHKAGVDPTEPFG 658

RESULT 10

Q6MCK4_PARUM PRELIMINARY; PRT; 456 AA.
 ID Q6MCK4_1
 AC Q6MCK4_1
 DT 05-JUL-2004 (Tremblrel. 27, last sequence update)
 DT 05-JUL-2004 (Tremblrel. 27, last annotation update)
 DE Hypothetical protein.
 GN OrderedlocusNames=pc0851;
 OS Parachlamydia sp. (strain UWE25) (subsp. Acanthamoeba sp.).
 OC Bacteria; Chlamydiae; Parachlamydiales; Parachlamydiaceae; Parachlamydia.
 NCBI_TaxID=264201;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RX PubMed=15073324;
 RA Horn M., Collingro A., Schmitz-Besser S., Beier C.L., Parkhoid U.,

RA NUCLEOTIDE SEQUENCE.
RC STRAIN=AVOP;
RG US DOE Joint Genome Institute (JGI-ORNL);
RA Larimer F., Land M.;
RT "Annotation of the draft genome assembly of *Azotobacter vinelandii*
AVOP";
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AVOP;
RA DOE Joint Genome Institute;
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
RN [4]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AVOP;
RG US DOE Joint Genome Institute (JGI-PGF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Detter C., Glavina T.,
RA Hammon N., Jarant S., Peticus S., Richardson P.;
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAU03000008; EMB05116.1; -; Genomic DNA.
SQ SEQUENCE 445 AA; 5117 MW; DFE646ACF0E73370 CRC64;

Query Match 23.3%; Score 816; DB 2; Length 445;
Best Local Similarity 43.2%; Pred. No. 9.3e-42;
Matches 202; Conservative 58; Mismatches 166; Indels 42; Gaps 15;

QY 231 LRRPAATA-----DTRRAVIDAP---GVPFHEVLALVGDARIVLIGESSHG 278
DB 1 MKRSDTSAAPETREHVELLRHAISLPALDSGFAS---LFRYRGDARVLLIGESHS 57
QY 279 THEFYQAANAATOMLIEKFGAVAAEADPDAVYVNVYVGLGDTNADALSGFEFRP 338
DB 58 TSEFRTPAATIRLRIEHEGNIYVAEADWPAADIDQVR---RQRPADBERQAFFQRP 114
QY 339 AMMNRTVVRDPEVEMLR--TNORYESGALRQAGFYGLDLSLHRSIOEVLSYLDKVDPR 396
DB 115 TMMRNLEVAEFQWLNRFNDLPYE---RVERRGIDVYSWASIGAVLDYDKTDQ 170
QY 397 AAAPARARYACFDHACDDGAYGFAAFAAGPSCEREAVQYLDVQNALAIVARQDGL 456
DB 171 AARAREYVGLT-MPWRDAPARYGDDVLYG--HOSCEBAVIEQLRSLLELRLEALGRD-- 226
QY 457 AEDLFAAQONAAQVYRADEVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYV 516
DB 227 --EDFFETQNAARVQAAEHYRTMYHSSASMNLRDHFDTLRHLEH---RQPEA 279
QY 517 RIVVMAHNSHVGDARATEV-WADGQLTLGQIVRERYGDESRISGFSTYTGVTAASEWCG 575
DB 280 RAVVMAHNSHIGMADATMGWA-GEFNIGELCRTMGDAVLIGTGTRGSAVACSEWDE 338
QY 576 IAQRKAVRPALHGVSEELFHQTA-DSFLVSNR--LSRDAEAPLDVLRGAIQVYVLPAT 632
DB 339 PMRIKQVPSRPSRDEWELFLOTGFPATLTDWRNDPSEILRAALSETLLERAIQVYVPS 398
QY 633 ERQSHYLVHRAPOFDMIHIDQFRALEPLEVTSRWIAGENPETYPPG 660
DB 399 ERQSHYFEALAEOPDAWIMLETHAVTPL--GPRQPOAGALETYPPFG 444

RESULT 13
OSBSS5_EMENT PRELIMINARY; PRT; 453 AA.
AC OSBSS5;
DT 10-MAY-2005 (TREMblrel. 30, Created)
DT 10-MAY-2005 (TREMblrel. 30, Last sequence update)
DT 10-MAY-2005 (TREMblrel. 30, Last annotation update)
DE Hypothetical protein.
GN ORNameB=AM4105.2;
OS *Aspergillus nidulans* FGSC A4.

OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiiales; Trichocomaceae; EmERICella.
OX NCBI_Taxid=227321;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=FGSC A4;
RA Birren B., Nisbaum C., Abouelleil A., Allen N., Anderson S.,
RA Arachchi H.M., Barna N., Baetien V., Bloom T., Boguslavsky L.,
RA Boukhalter B., Butler J., Calvo S.E., Camarata J., Chang Y.,
RA Choepel Y., Collymore A., Cook A., Cooke P., Corum B., Deatellano K.,
RA Diaz J.S., Dodge S., Dooley K., Dorris L., Elkins R., Engels R.,
RA Erickson J., Faro S., Ferreira P., Fitzgerald M., Gage D., Galagan J.,
RA Gardyna S., Gnetre S., Graham L., Grand-Pierre N., Hafez N.,
RA Hagopian D., Hago B., Hall J., Horton L., Hulme W., Iliev I.,
RA Jaffe D., Johnson R., Jones C., Kamal M., Kamat A., Karatae A.,
RA Kells C., Landers T., Levine R., Lindblad-Toh K., Liu G., Lui A.,
RA Ma L.-J., Mabbitt R., Maclean C., Macdonald P., Major J., Manning J.,
RA Matthews C., Mauceli E., McCarthy M., Meldrum J., Meneus L.,
RA Milnova T., Mlenga V., Murphy T., Naylor J., Nguyen C., Nicol R.,
RA Nielsen C.B., Norbu C., O'Connor T., O'Donnell P., O'Neill D.,
RA Oliver J., Peterson K., Phunhkhong P., Pierre N., Purcell S.,
RA Rachupka A., Ramasamy U., Raymond C., Retta R., Rise C., Rogov P.,
RA Roman J., Schauer S., Schupbach R., Seaman S., Severy P., Smitnov S.,
RA Smith C., Spencer B., Stange-Thomann N., Stojanovic N., Stubbs M.,
RA Talmas J., Testaye S., Theodore J., Topham K., Travers M.,
RA Vassiliev H., Venkataraman V.S., Viel R., Vo A., Wang S., Wilson B.,
RA Wu X., Wyman D., Young G., Zaitoun J., Zembek L., Zimmer A., Zody M.,
RA Lander E.;
RL "Genome Sequence of *Aspergillus nidulans*";
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AACD01000067; EAA59366.1; -; Genomic DNA.
KM Hypothetical protein.
SQ SEQUENCE 453 AA; 51811 MW; B9B90F5B67C4A21E CRC64;

Query Match 23.3%; Score 816; DB 2; Length 453;
Best Local Similarity 43.5%; Pred. No. 9.5e-42;
Matches 188; Conservative 58; Mismatches 156; Indels 30; Gaps 12;

QY 268 RIVLIGSSHGCTHEFYQAANAATOMLIEKFGAVAAEADPDAVYVNVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYV 320
DB 34 RIVLIGSSHGCTHEFYQAANAATOMLIEKFGAVAAEADPDAVYVNVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYV 93
QY 321 LGE---DTNADALSGFERPPAMMRTVVRDPEVEMLR--TNORYESGALRQAGFYGLDLY 377
DB 94 IGETKGADYEDGEAFERFTYMMKRNKEMQDLEWYRDINAFLP--AEKKVGIYGMIDLY 151
QY 378 SLHRSIOEVLSYLDKVDPRRAAPARARYACFDHACDDGAYGFAAFAAGPSCEREAV 437
DB 152 SLGSMKALISYLDKVDPRRAAPARARYACFDHACDDGAYGFAAFAAGPSCEREAV 209
QY 438 QLYDVQONALAAVYARQDGLAEDLFAAQONAAQVYRADEVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYV 497
DB 210 MLRDLLERRRLKYAHNP--DGEFHSSEQNAAYVRAERYKKMYVSSAGSWTLRSHMV 267
QY 498 QTLGSLTLHLDRLHLDAPPARIVVMAHNSHVGDARATEVMA--DQQLTLGQIVRERYGDESR 556
DB 268 DTLKRLIKH-----KPDGKVVYVMAHNSHCGDARHTRMGRRREVNIQGLCREVPEQNV 321
QY 557 S-IGFSTYTGVTAASEWGGIAQRKAVRPALHGVSEELFHQTA-ADSEFLVSA---RLSRDA 611
DB 322 ALIGCGTHGTVAASAASWGEDWQIMVDPKPSRDSWRIADHTGVESFLDLRDERSREV 381
QY 612 EAPL--DVVLRGALIGVYVLPATERQSHYLVHRAPOFDMIHIDQFRALEPLEVTSRWI 669
DB 382 REALAKKHRLQRFIGIYVPRTERISHVSAADLVNQFDGVVDPVTTSAVPLEVQPKT 441
QY 670 AGENPETYPPGL 681
DB 442 ALGDERTYPPGL 453

```
RESULT 14
Q4WJ15 ASPFU PRELIMINARY; PRT; 463 AA.
ID Q4WJ15
AC Q4WJ15
DT 13-SEP-2005 (TReMBLrel. 31, Last Created)
DT 13-SEP-2005 (TReMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TReMBLrel. 31, Last annotation update)
DE Erythromycin esterase family protein.
GN ORFNames=Afu1G05850;
OS Aspergillus fumigatus Af293.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=330879;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Af293;
RA Niernan W., Pain A., Anderson M.J., Wortman J., Kim H., Stanley,
RA Artaya J., Britman M., Abe K., Archer D.B., Bermejo C., Bennett J.,
RA Bowyer P., Chen D., Collins M., Coulson R., Davies R., Dyer P.S.,
RA Farman M., Fedorova N., Fedorova N., Feldblyum T.V., Fischer R.,
RA Foker N., Fraser A., Garcia J.L., Garcia M.J., Goble A.,
RA Goldman G.H., Gomi K., Griffith-Jones S., Gillman R., Haas B.,
RA Haas H., Harris D., Horiuchi H., Huang J., Humphrey S., Jimenez J.,
RA Keller N., Knout H., Kitanoto K., Kobayashi T., Kulkarni R.,
RA Kumagai T., Lafton A., Latge J.-P., Li W., Lord A., Lu C.,
RA Majoros W.H., May G.S., Miller B.L., Mohamoud Y., Molina M., Monod M.,
RA Moulyna I., Mulligan S., Murphy L., O'Neill S., Paulsen I.,
RA Penalta M.A., Perteu M., Price C., Pritchard B.L., Quail M.A.,
RA Rabinowitsch E., Rawlins N., Rajandream M.-A., Reichard U.,
RA Renauld C.M., Rutter S., Salzberg S.L., Sanchez M.,
RA Sanchez-Ferrero J.C., Saunders D., Seeger K., Squares R., Squares S.,
RA Takeuchi M., Tekra F., Turner G., Vazquez de Aldana C.R., Weidman J.,
RA White O., Woodward J., Yu J.-H., Fraser C., Galagan J.E., Asai K.,
RA Machida M., Hall N., Barrett B., Denning D.W.;
RT "Genomic sequence of the pathogenic and allergenic filamentous fungus
RT Aspergillus fumigatus";
RL Submitted (May-2005) to the EMBL/Genbank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/Genbank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC EMBL; AAF0100007; EAL88297.1; -; Genomic DNA.
DR EMBL; AAF0100007; CB41E1508472202 CRC64;
SQ SEQUENCE 463 AA; 53158 MW; CB41E1508472202 CRC64;

Query March 23.1%; Score 811.5; DB 2; Length 463;
Best Local Similarity 42.0%; Pred. No. 1.8e-41;
Matches 198; Conservative 66; Mismatches 172; Indels 35; Gaps 16;

QY 230 SLRRPASTADVADRVVADAPG-GVPTHEVLAELVGDARIIVLIGESSHGHTEFYQARA 288
DB 9 SSRPMAAOLQOLTEAAQPIPSIDBSFGSHFNFNGYKVVLLGDSHGSEFYAARAE 68
QY 289 MTQWLIBEKGFAVAEAADWDPAAYRVNRYV-----GLGEDTNADALSGFERPAM 341
DB 69 ITRKLIBQHGTYWALAEADWDPAEDRIDRYVRPGRPKAGIC---GKANDLEPFKRFPTM 125
QY 342 WRNTVVDVFWMLTRNQRYSAGLRQAGFYGLDLSLRHSIQEIVSTLDKVDPRRAARA 401
DB 126 WRNEMODLVEMNRNRAKLPNE--KAGVYGLDLSGASIRAVITDLDVDPAGKEA 183
QY 402 RARACFPDHCADQGAAGFAAGAGPSCEBAVEOLVDVORNALAARODGLAEBDEL 461
DB 184 RRRIGCL-QPWNDEBSAYGLASLRGM--EDCESGVLTQMLRDLKRLQYAGND--VRGDE 239
QY 462 FYA-QQNAQTVRDAEYVYRAMFSGRVTSWNLRDQMAQTGLSLTHLDRLHADPARIV 520
DB 240 FHSQQAQNAFVVRDAERYKAYYSASWSLTDTHMFTTLRLFRH-----KPAQAKIV 294
QY 521 WAHNSHVDAAATERYMA-DGQTLTGQIYRERYGDSRSI-GFSYTGTVTAASWGGIAQ 578
DB 295 WAHNSHCGDAAYTSGTRNREVNIGQLIRENFGRBNALVILGCGHTGTVAHAHEMDDMQ 354
```

```
QY 579 RKAVRPLHGSVEFLPQTA-DSFLVSARLS-----RDEAPLDVVRIGRAIGVYLP 631
DB 355 VMKRPSPRSDSWETIADTGISFVVIDLRKHLDPALRTMA-AENSRLRFIVGIVRP 413
QY 632 TEROSHYLAVRPADQPDAMIHDQTRALBPLE-VTSRWIAGENPETYPTGL 681
DB 414 TERISHYSQALYNQPDAYIWPDTVAENKPLEKVPKPTPLGQP-ETYPFGV 463

RESULT 15
Q41713 GIBZE PRELIMINARY; PRT; 455 AA.
ID Q41713
AC Q41713
DT 13-SEP-2005 (TReMBLrel. 31, Created)
DT 13-SEP-2005 (TReMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TReMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=FG06995.1;
OS Gibberella zeae PH-1.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
OX NCBI_TaxID=229533;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PH-1;
RA Birren B., Nusbaum C., Abouelleil A., Allen N., Anderson S.,
RA Archchi H.M., Barua N., Bastien V., Bloom T., Bogulavsky L.,
RA Boukhalter B., Butler J., Calvo S.E., Camarata J., Chang Y.,
RA Choepel Y., Collymore A., Cook A., Cooke P., Corum B., Dearellano K.,
RA Diaz J.S., Dodge S., Dooley K., Dorris L., Elkins T., Engels R.,
RA Erickson J., Fato S., Ferreira P., FitzGerald M., Gage D., Galagan J.,
RA Gadyana S., Ghebre S., Graham L., Grand-Pierre N., Hafez N.,
RA Hagopian D., Hagos B., Hall J., Horton L., Hulme W., Iliev I.,
RA Jaffe D., Johnson R., Jones C., Kamal M., Kamat A., Karatae A.,
RA Kells C., Landers T., Levine R., Lindblad-Toh K., Liu G., Lui A.,
RA Ma L.-J., Mabbitt R., Maclean C., Macdonald P., Major J., Manning J.,
RA Matthews C., Maucelli E., McCarthy M., Melidri J., Menues L.,
RA Milnova T., Mienga V., Murphy T., Naylor J., Nguyen C., Nicol R.,
RA Nielsen C.B., Norbu C., O'Connor T., O'Donnell P., O'Neill D.,
RA Oliver J., Peterson K., Phunkhang P., Pierre N., Purcell S.,
RA Rachupka A., Ramasamy U., Raymond C., Retta R., Rise C., Rogov P.,
RA Roman J., Schauer S., Schupbach R., Seaman S., Severy P., Smitov S.,
RA Smith C., Spencer B., Stange-Thomann N., Stojanovic N., Stubbs M.,
RA Talamas J., Tesfaye S., Theodore J., Topham K., Travers M.,
RA Vassiliev H., Venkataraman V.S., Viel R., Vo A., Wang S., Wilson B.,
RA Wu X., Wyman D., Young G., Zainoun J., Zembek L., Zimmer A., Zody M.,
RA Lander E.;
RT "Fusarium graminearum genome sequence";
RL Submitted (Feb-2004) to the EMBL/Genbank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/Genbank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC EMBL; AAC0100289; EAA76190.1; -; Genomic DNA.
DR EMBL; AAC0100289; EAA76190.1; -; Genomic DNA.
KW Hypothetical protein.
SQ SEQUENCE 455 AA; 51820 MW; C3148AAE899A6D7F CRC64;

Query March 22.8%; Score 800.5; DB 2; Length 455;
Best Local Similarity 42.4%; Pred. No. 8.5e-41;
Matches 185; Conservative 62; Mismatches 156; Indels 33; Gaps 12;

QY 265 GDARIVLIGSSHGHTEFYQARAAMTQWLIBEKGFAVAEAADWDPAAYRVNRYV----- 319
DB 34 GDCVLLIGDASHGTSFYSVRAEITKVTMIENGNFINVAEADWSDAEHVDYVRHRPVP 93
QY 320 --GLGEDT--NADALSGFERPAMWRRTTVTRDPEWMLRTNQRYESG--ALRQAGFY 372
DB 94 GQGAVERITQWAEKEKESPLRPTWWRNVEVHDPEWRS-----YNSGREVTEAAGFY 149
QY 373 GLDYSLSHRISQEVISTYLDVDPRAAARARYA-CFDHACADQGAAGFAAGAGPSC 431
DB 150 GLDYSLSGTSKMAVIDYLDVDPQAKVARGRIYNLMD--WLEDPEHYGLSGLATSKGY 207
QY 432 EREAVEQLVDVORNALAY-ARODGLAEBDELFYAQQNAQTVRDAEYVYRAMFSGRVTSWN 490
```

```

Db      208 EQDVVAMIGDLLRRKRIEYSAALDGV---EFHNGEQNARVVKDAEQYKAMYRGODKSWN 264
QY      491 LRQDMQOTLGSLLTHLDRHLDAPPARIVWVWNSHVGDARATEY-WADGOLTLGOIVRE 549
Db      265 HRDMHMFETLKRVLH-----RGEQSKAIWVWNSHIGDADATSMNSHSHLNGELCKR 319
QY      550 RYGDERSRISGFSTYTGVTAASEWQIAQRKAVRPALHGSVEELFHQTA-DSPVLSARLS 608
Db      320 AFGDHALSIGTGTWGTVAQAQNWESDWNIIKVQPLPGSYEELMHATGIGNFVLDLRKG 379
QY      609 RDAEAPLDV--RLGRAIGVYVLPATERQSHYLVPRADQPDAMIHIDQTRALEPLEVT 665
Db      380 KCDEKLREALNGERLERSIGVIYKPEYKASHYSYAILPDQDGYWPEDESKHVGTLLEIH 439
QY      666 SRWLAGENPETYPQGL 681
Db      440 QPRSPLEYHETWPPGL 455

```

Search completed: March 23, 2006, 05:24:20
 Job time : 177.174 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 23, 2006, 05:24:53 ; Search time 46.4465 Seconds
(without alignment)

1212.194 Million cell updates/sec

Title: US-10-617-038-23

Perfect score: 3508
Sequence: 1 VLMTAADVTRSPRRVFRD.....LEVTSRWIAGENPEPTPTGL 681

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/1/1aa/5.COMB.pep.*
2: /cgn2_6/ptodata/1/1aa/6.COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/H.COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/PCUTS.COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/RB.COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	547.5	15.6	226	US-09-902-540-10451	Sequence 10451, A
2	417.5	11.9	229	US-09-902-540-15374	Sequence 15374, A
3	415.5	11.8	229	US-08-311-731A-396	Sequence 396, App
4	346.5	9.9	227	US-09-902-540-12846	Sequence 12846, A
5	147.5	4.2	1650	US-09-252-991A-21798	Sequence 21798, A
6	143.5	4.1	2214	US-09-902-540-15988	Sequence 15988, A
7	136	3.9	685	US-09-252-991A-26723	Sequence 26723, A
8	134.5	3.8	3031	US-07-689-008-2	Sequence 2, Appli
9	134	3.8	3562	US-09-679-279-14	Sequence 14, Appli
10	131	3.7	535	US-09-252-991A-17140	Sequence 17140, A
11	131	3.7	1275	US-08-426-630-49	Sequence 49, Appl
12	130	3.7	6396	US-09-410-551B-72	Sequence 72, Appl
13	130	3.7	6396	US-09-940-316B-72	Sequence 72, Appl
14	128.5	3.7	2841	US-09-477-962-117	Sequence 117, Appl
15	128.5	3.7	3413	US-10-042-665A-8	Sequence 8, Appli
16	128	3.6	923	US-09-252-991A-22409	Sequence 22409, A
17	127.5	3.6	7257	US-09-325-409-5	Sequence 5, Appli
18	127.5	3.6	7257	US-09-568-102-5	Sequence 5, Appli
19	127.5	3.6	7257	US-09-567-969-5	Sequence 5, Appli
20	127.5	3.6	7257	US-09-568-480-5	Sequence 5, Appli
21	127.5	3.6	7257	US-09-568-486-5	Sequence 5, Appli
22	127.5	3.6	7257	US-09-568-472-5	Sequence 5, Appli
23	127.5	3.6	7257	US-09-567-899-5	Sequence 5, Appli
24	127.5	3.6	7257	US-10-014-717-5	Sequence 5, Appli
25	127	3.6	1411	US-09-252-991A-23628	Sequence 23628, A
26	126.5	3.6	903	US-09-902-540-11595	Sequence 11595, A
27	126.5	3.6	3739	US-09-320-878-2	Sequence 2, Appli

28	126.5	3.6	3739	2	US-09-105-537-33	Sequence 33, Appl
29	126.5	3.6	3739	2	US-09-141-908-3	Sequence 3, Appli
30	126.5	3.6	3739	2	US-09-657-440-2	Sequence 2, Appli
31	126.5	3.6	3739	2	US-09-793-708-2	Sequence 2, Appli
32	126.5	3.6	11877	2	US-09-105-537-6	Sequence 6, Appli
33	126	3.6	577	2	US-09-902-540-14076	Sequence 14076, A
34	125.5	3.6	1053	2	US-09-902-540-12126	Sequence 12126, A
35	125	3.6	852	2	US-09-902-540-13582	Sequence 13582, A
36	124.5	3.5	2108	2	US-09-252-991A-31502	Sequence 31502, A
37	124	3.5	1875	2	US-10-042-665A-2	Sequence 2, Appli
38	124	3.5	1939	2	US-10-152-886-43	Sequence 43, Appl
39	123	3.5	879	2	US-09-252-991A-27857	Sequence 27857, A
40	123	3.5	879	2	US-09-252-991A-22604	Sequence 22604, A
41	123	3.5	4472	2	US-08-804-227C-2	Sequence 2, Appli
42	122	3.5	729	2	US-09-252-991A-32535	Sequence 32535, A
43	122	3.5	765	2	US-09-252-991A-19778	Sequence 19778, A
44	122	3.5	831	2	US-09-252-991A-30097	Sequence 30097, A
45	121	3.4	1891	1	US-08-804-227C-12	Sequence 12, Appl

ALIGNMENTS

```
RESULT 1
US-09-902-540-10451
; Sequence 10451, Application US/0902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(11549)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 10451
; LENGTH: 226
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-10451

Query Match      15.6%; Score 547.5; DB 2; Length 226;
Best Local Similarity 51.8%; Pred. No. 3.9e+46;
Matches 118; Conservative 30; Mismatches 73; Indels 7; Gaps 2;

QY 12 RSPRRVFRDRRAGRVLAELAAVROPDIVIGLAGGLPVAVEVAALHAPLDAFVR 71
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 2 REPV--FQDRYEGRGVLAADHRRARPGTIVLALPRGQPVGVAGVARKGVPLDVLVR 59
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 72 KLAGPGRHDEPAVAGALASGGRVVNDVVRGLRITPQQLRDIAREGRELRLRESAYRGER 131
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 60 KLGPGHGEELAMGAIASGGRVLRVREVNINIGREQIEAARREAVELQRRRESYRGER 119
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 132 PPTITGKTIVYVNDGLATGASMPAAVOALRDQAPQIVAVAPPAPESTCRFAGLVDDV 191
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 120 PPPVRRRTVTLVDDGLATGTMPAAVALRLLEPAIIVAVPVAALSCSGLQVADEV 179
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 192 VCAMPPTPLAVGESFMDPQVTDVEVRRLATPTAAPSRLRRPASTA 239
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 180 IGVMPPEFFSVGLMYRDPQTEDEVRALLAEXT-----REGAASRA 222
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 2
US-09-902-540-15374
; Sequence 15374, Application US/0902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
```

APPLICANT: Slater, Steven C.
APPLICANT: Wiegand, Roger C.
TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
FILE REFERENCE: 38-10(15849)B
CURRENT APPLICATION NUMBER: US/09/902,540
CURRENT FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: 60/217,883
PRIOR FILING DATE: 2000-07-10
NUMBER OF SEQ ID NOS: 16825
SEQ ID NO 15374
LENGTH: 229
TYPE: PRT
ORGANISM: Myxococcus xanthus
US-09-902-540-15374

Query Match
Best Local Similarity 42.5%; Score 417.5; DB 2; Length 229;
Matches 96; Conservative 39; Mismatches 86; Indels 5; Gaps 3;

QY 18 FRDRRAGRVLAELL--AAVRDQPDVIVLGLARGGLPVAMEVAAALHAPLDAFVVRKLG 75
DB 3 FEDRVDAAGRRLAQLLRGGYTG-GLIVLALPRGGVPVAFVATALGAPLDVWVRKGV 61
QY 76 PGHDEFVAGALASGRVYVNDVVRGLRITPQQLRDIAERGRLLRRESAYRGERPPTD 135
DB 62 PGYELGIGAVAEQVAFVNRRLMDEVGVTEDMQGLVRQKTDVKRVAFRQGIABPR 121
QY 136 ITGKTVIVDDGLATGASMFVAVOALRDAPQIVIAVPAPESTCRFAGLVDDV 195
DB 122 IEGRRITLVDDGATGCGVRAAIOALNRHRSIIIVAPVAASQTLAEPLADVV 181
QY 196 MPTPLAVGESFMDFRQVTDSEVRRLATP--TAGPSLRPAASTA 239
DB 182 STPLVAIGQWYADFQVDPDAVATLLAQLRLTLGHDRHAPRGPDTA 227

RESULT 3
US-08-311-731A-396
Sequence 396, Application US/08311731A
Patent No. 6583266
GENERAL INFORMATION:
APPLICANT: SMITH, DOUGLAS
APPLICANT: MAO, JEN-I
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
TITLE OF INVENTION: RELATING TO MYCOBACTERIUM TUBERCULOSIS AND LAPRAE FOR
TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 411
CORRESPONDENCE ADDRESS:
ADDRESSEE: WOLF, GREENFIELD & SACKS, P.C.
STREET: 600 ATLANTIC AVENUE
CITY: BOSTON
STATE: MASSACHUSETTS
COUNTRY: USA
ZIP: 02210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/311,731A
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: GATES, EDWARD R.
REGISTRATION NUMBER: 31,616
REFERENCE/DOCKET NUMBER: C0044/7125
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/720-3500
TELEFAX: 617/720-2441
INFORMATION FOR SEQ ID NO: 396:
SEQUENCE CHARACTERISTICS:
LENGTH: 289 amino acids

TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Mycobacterium leprae
US-08-311-731A-396

Query Match
Best Local Similarity 41.6%; Score 415.5; DB 2; Length 289;
Matches 97; Conservative 33; Mismatches 90; Indels 13; Gaps 3;

QY 12 RSPRRVFRDRRAGRVLAELLAAVRDQPDVIVLGLARGGLPVAMEVAAALHAPLDAFVVR 71
DB 53 RNSMRLLFDDRVDAAGRRLAQLLRGGYTG-GLIVLALPRGGVPVAFVATALGAPLDVWVR 111
QY 72 KLGAPGHDEFVAGALASGRVYVNDVVRGLRITPQQLRDIAERGRLLRRESAYRGERPPTD 131
DB 112 KLGAPGHDEFVAGALASGRVYVNDVVRGLRITPQQLRDIAERGRLLRRESAYRGERPPTD 171
QY 132 PPTDITGKTVIVDDGLATGASMFVAVOALRDAPQIVIAVPAPESTCRFAGLVDDV 191
DB 172 DRAPIGGRITVIVDDGATGCGVRAAIOALNRHRSIIIVAPVAASQTLAEPLADVV 231
QY 192 VCATMPTPLAVGESFMDFRQVTDSEVRRLATP--TAGPSLRPAASTA 232
DB 232 VCLHTPAFCVAVQGYCNFTQSDAEVIALDRAREGFESATTATLGPQIR 284

RESULT 4
US-09-902-540-12846
Sequence 12846, Application US/09902540
Patent No. 6833447
GENERAL INFORMATION:
APPLICANT: Goldman, Barry S.
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Wiegand, Roger C.
TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
FILE REFERENCE: 38-10(15849)B
CURRENT APPLICATION NUMBER: US/09/902,540
CURRENT FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: 60/217,883
PRIOR FILING DATE: 2000-07-10
NUMBER OF SEQ ID NOS: 16825
SEQ ID NO 12846
LENGTH: 227
TYPE: PRT
ORGANISM: Myxococcus xanthus
US-09-902-540-12846

Query Match
Best Local Similarity 9.9%; Score 346.5; DB 2; Length 227;
Matches 84; Conservative 34; Mismatches 84; Indels 3; Gaps 3;

QY 18 FRDRRAGRVLAELLAAVRDQPDVIVLGLARGGLPVAMEVAAALHAPLDAFVVRKLGABG 77
DB 2 FRDRRAGRVLAELLAAVRDQPDVIVLGLARGGLPVAMEVAAALHAPLDAFVVRKLGABG 60
QY 78 HDEFVAGALASGRVYVNDVVRGLRITPQQLRDIAERGRLLRRESAYRGERPPTDIT 137
DB 61 R-WTVLGAIVSEGGIYLDQDALRGGLPEVARSILARASAEVDNQVRLRGTSEPA-WG 118
QY 138 GKTIVIVDDGLATGASMFVAVOALRDAPQIVIAVPAPESTCRFAGLVDDV 197
DB 119 GFTVLVDDGLVSGATAMALDVLRRQHPARLVGVSTPHGLARVREDAVHCV 178
QY 198 TPPLAVGESFMDFRQVTDSEVRRL 222
DB 179 PAMRDVSEAYDVRPLVDVLRQL 203

RESULT 5

US-09-252-991A-21798
; Sequence 21798, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 21798
; LENGTH: 1650
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-21798

Query Match 4.2%; Score 147.5; DB 2; Length 1650;
Best Local Similarity 22.1%; Pred. No. 3.7e-05;
Matches 172; Conservative 75; Mismatches 304; Indels 227; Gaps 36;

QY 43 VTGLARGGLPVAMEVVAALHAPLDAF-----VVRKLGAPGHDFEFAVGA-----LAS 88
DB 269 VTSLSPEFMFALSGAGVH-PLRYLADPGALAVVQAAEGDGLAPCSKGRRCRCP 327
QY 89 GGRVVVNDVVRGLRITPQQLRDIAEREGRELLRESAYRGERPPTDTGKTVIVDDGL 148
DB 328 GSRLV-----TVRPMPSGQAGVAVRGV-----VQAQFPGADDAPLLAVDADGQ 374
QY 149 ---ATGASMPAAVQALRDAOP-----AQIVAVPA----- 175
DB 375 QLEALLAGNLAALAAVEDVQPLDAGRAASTDQATLVQVAAPOVETDAALADQPALHQ 434
QY 176 -----AP-----ESTCREFAGLVDDVVCATMPTPFLAVGESFMDPROVTDDE 217
DB 435 ALDQAQVAFPGDPAAVGVDAFCGQF-----QFAVAGQAATVQVQGAQ 480
QY 218 VRRLATPTAGPSLRPAAS---TAAD-----VLRRVALDAPGVPPTHVLAELV-GDA 267
DB 481 AQRLLATGGAATVVEGAAVQVHMLADQATLVGVQASGDAQAGVAGEQALAAVQCGA 540
QY 268 RIVLIGE---SSSGTHEFYQARAAMTQMLIEKGFQVAA-----EADNPDARV 314
DB 541 -----GERQAGAGNGAFGVVQRAVDAQVDAAGQAGALGVVQASGRDVQAGAGDHG 595
QY 315 NRYVRIGGEDT---NADEALSGFERPRAMMMRNTVVADVFWELTRNQRYESGALRQAG 370
DB 596 QAVVQGLDADQAGQAGADQAPAAVAVQAGSGEGEAGDFPRAVAV--HRELAAQQQAG 653
QY 371 FYGLDLYSLHRSIOEVI SYLDKVDPRAAARARAYACFDHACADDDQAYGFPAAGAPS 430
DB 654 -----RGDQAAVA-----VDQRAAIYVEG-----DAGPADDAASALVEAFQV-- 692
QY 431 CERAEVQVLDVQVNNALAYARODGILLAEDELFYAQONNAQTVRDAAVYTRAMPSGRVTSWN 490
DB 693 -----VEQALDAGSLLAVQPG-----HGGDAGVADAADAVALVQAHAG----- 732
QY 491 LRDMHAGTGLSLTH-----LDRH---LDAPRARIYV-----WAINSHVGDARA 532
DB 733 -ADH--RITGA--DHAGTAVVEAGALQRIHAGIAEQRIALVVQKGLAQKQRTGAGEGPA 787
QY 533 TEVWADGQLTIGQIVRER-----YGDERSRISGFSTYTGTVTAASEWGIQORKAVERPAL 586
DB 788 TVVQARGARQQAFAFDQRAALVVQAAABAHQAQVLAVERPAVAVQFPAV--QAQAQVAAQ 846
QY 587 H--GSVEELFHQIADSF---LVSARLSRDAEAPLDVVRVLRGAIG-VYUPLATERQSHYL 639
DB 847 HPLGLVQQAALHGEQAQAAVADDLAAAVVQLLAGVGHDLRGAGNLAGAVALDPLRLSDAAR 906

QY 640 HVRP-----ADQFDMIHIDQTALEBLEVTSWIAGENPETVPTGL 681
DB 907 GDQGLAVVDRVGRDLQGLFADQCFATLLGQAARRLQVA-----LGGDTPSGVADGV 958

RESULT 6
US-09-902-540-15988
; Sequence 15988, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(115849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 15988
; LENGTH: 2214
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-15988

Query Match 4.1%; Score 143.5; DB 2; Length 2214;
Best Local Similarity 23.1%; Pred. No. 0.00016;
Matches 173; Conservative 80; Mismatches 302; Indels 193; Gaps 37;

QY 5 AADVTRSPRRVFRDRREAGRVLAELAAVRDQ-PVIVGLARGGLPVAMEVVAALHA 63
DB 1278 ATAELVYRRADLEDDKDA---IDALLPLARLRPADASVIDRA-----VAGLHA 1324
QY 64 PLDAFVVRKLGAPGH--DEPAVGLASGGRVVNDVVRGLRITPQQLRDIAEREGRELL 121
DB 1325 -----LGRHDLDDVVAASGAEAAGRRAA-ELLAAASVASSST---ADEDDAWTL 1371
QY 122 RRESAYRGERPPTDTGKTVIVDDGLATGASMPAAVQALRDAQVQIVAVPAPESTC 181
DB 1372 TORAA---EAPBEDTLALQALVT--GLRQKDSFRLLEALE---QWPRVEDADEAV 1421
QY 182 REF-----AGVDDVVCATMPTPFLAVGESFMDPROVTDSEVRRLATPTAGPSLR 233
DB 1422 LRLELAELADAGVD-----VARELDEVAARAGAGAGAG-- 1456
QY 234 PAASTADVLRRAVIDAPGVPTHEVL-ABELV-GDARIVLIGSSHGTHEFYQARAAMTQ 291
DB 1457 ---ADALEALEPLLKDAPARAEVQVABELASGRBOVLLMAAAG---FESAGQLPE 1509
QY 292 WLIEKGFQVAAADMPDAVRVNRVYRGLGEDTNADEALSGFERPRAMMMRNTVVADRV 351
DB 1510 ALKRAAK--DAASVPDVALRVNHLVRSAGEAPRAARALQARLARLAPEERPLE-- 1565
QY 352 EWLTRNQRYESGALRQAGFYGLDLYSLHRSIOEVI SYLDKVDPRAAARARAYACFDHA 411
DB 1566 -----AAGIMEXAGEHEALYVFERIATA---SDVLABSELAEKRGRLGAPARA 1612
QY 412 CADDQAYGFPAAGAPGSPSCREAVEQLDVQVNNALAYARQD-----LLAE 458
DB 1613 L-----EVGFAPAAASG-----DLTD---ALMAAQAGDTARTREALMALALAD 1654
QY 459 DELFYAQONNAQTV--DAE--VYTRAMPSGRVTSW--NLRDQMAQTLGSLTHLDRHL 511
DB 1655 ADPNAASALDGLRAEBDABGLELELAGISAARDAAFAFVALRDEVLRSAASVLRRL-RAL 1713
QY 512 D--APP-----ARIYVMAHNSHVGDAR-----ATEVWADGQLT 542
DB 1714 BELAPBEGFAARLTLPLTEKLPBALAEAVLTQVNAQPGIVRBEALMAADGSPSRKS 1773
QY 543 LGOIVRERRYGDERSRISGFSTYTGTVTAASEWGIQORKAVERPALGSEVEL---FHQTA 598

Db 1774 L---LRRRYALE---LELGRYACVRLTSLQVAGESSDAVANAVALHLERGBLLSLPLQLP 1827
QY 599 DSFLVSRARLSDAPAE-----LDVYRLGALGVYVLPATERQSHLYLV-----RP 643
Db 1828 EAEAEAFQALKDPAQQLHSRLHLLSLVDLGRH-PAVFLSVLERLEQAGVSGNTYERRL 1886
QY 644 ADQFDM-IHIDPRLPLEVTSRWTA 670
Db 1887 ADAVEALGVADAQAQLERLPLETERLTA 1914

RESULT 7

US-09-252-991A-26723
; Sequence 26723, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 26723
; LENGTH: 685
; TYPE: PRN
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26723

Query Match 3.9%; Score 136; DB 2; Length 685;
Best Local Similarity 24.9%; Pred. No. 0.00013;

Matches 133; Conservative 57; Mismatches 177; Indels 164; Gaps 34;

QY 9 VTRRSRRVRRDRRAGRVLAELLAAYRDPDVLVGLARGGLPVAME-----VAAAL 61
Db 108 VQRR--RRFAGAPRQAGROLAHA-----RGAVDVVDVGTVAESAQFAYGGGQGGMAADL 161
QY 62 HARLDAFVVRKLGAP-----GHDFPAGVALASGVVNVDDVVRGRTTPQOLR-DIA 113
Db 162 E---EAVVQAEVGMFQHFLEGGHQLQFAGGHA-----LVESFLALVVQVSTQSLAVDLA 214
QY 114 ERBGRRLRRRSAY-----RGRPRPTDITKTVIVVDG-----LATGA- 152
Db 215 VDRGRQAVQGVAAQGVVQGFLLAQGGDGEHGR-----GIAVLILADGHEGAQLLATGAW 270
QY 153 ---SMPAAVQALRDAPQAVIIVAPPAPESTCEEPAGLVDDVVCATMPFPFLAVGESFW 208
Db 271 VGRKHHRRAVDS-RLAQQRGLDVA-----EFPAVADLHLVVQPT----- 309
QY 209 DFRQVNDDEVRRLATATPAGSLRRPASTADVLRVAIDAPGG-VPTHEVLAEIVGDA 267
Db 310 -----EELQALIGQPA---GLVAGAVEPTPQCLR--VEAPRGHVRVAQVFAK-QADA 356
QY 268 RIVLIGSSHGHTHEFYQARAAMTQWLIEKGFAGVAAEAD---WPDAYRVVRVY---RG 320
Db 357 RQVQL-----TGSVMAARPVVG---VEDECLAVDDRADRGSGPMEFGVDDAAGSDRG 407
QY 321 LGEEDTNADEL---SGFER---FPAMWKRVTVADFEVEMLTRNQRYESGALRQAGFYGL 374
Db 408 LGRVAVVTHHGIGQPCGDRQFFPG-----AEQVAQR-QRRRGGLLRGRH--- 452
QY 375 DIVSLHSIGSEVSYLDKVDPRRAARARVACFDH-----ACADGGQAV--GFAA 423
Db 453 -----HNGRREGVA---DPLGEECCQRRRVAAHLVGHQVEAGACAGGPPDFLGNVE 502
QY 424 A---FGAGPSCREBA-----VEQLVDVQRNALAVARODGLIADEL 461
Db 503 ADAGHOGGAARPLIHIEARVVRPAHEVGQATLDHNLALGFA--GGARGEDHI 550

RESULT 8

US-07-689-008-2
; Sequence 2, Application US/07689008
; Patent No. 5268274
; GENERAL INFORMATION:
; APPLICANT: Ben-Bassat, Arle
; APPLICANT: Calhoun, Roger D
; APPLICANT: Fear, Anna L
; APPLICANT: Gelfand, David H
; APPLICANT: Meade, James H
; APPLICANT: Tal, Rony
; APPLICANT: Wong, Hing
; APPLICANT: Benizman, Moshe
; TITLE OF INVENTION: METHODS AND NUCLEIC ACID SEQUENCES FOR THE
; TITLE OF INVENTION: EXPRESSION OF CELLULOSE SYNTHASE OPERON
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: McCutchen, Doyle, Brown & Eneersen
; STREET: Three Embarcadero Center
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/689,008
; FILING DATE: 19910422
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 337,194
; FILING DATE: 12-APR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 496,236
; FILING DATE: 23-MAR-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Murphy, Lisabeth Felix
; REGISTRATION NUMBER: 31547
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 393-2000
; TELEFAX: (415) 393-2286
; TELEX: 340817 MACPAG SFO
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3031 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-689-008-2

Query Match 3.8%; Score 134.5; DB 1; Length 3031;
Best Local Similarity 19.5%; Pred. No. 0.0023;

Matches 152; Conservative 96; Mismatches 298; Indels 235; Gaps 34;

QY 16 RVFRDRRERAGRVLAEL-----LAAYRDPDVLVGLARGGLPVAMEVAAALHARLDAF 68
Db 1465 QILAQKKNQAKTIGDVLNAGDDLTSYRSSPLTYV-----GYFPLWLEPRVWVHNNPSRV 1519
QY 69 VVRKLAPGHDFPAGVALASGGRVNVDDVVRGLRTTPQQLDIAREBGRRLRRSAYR 128
Db 1520 IV-----VGLL---GCILIVAVVVRAL-----AGHALRRRRELQ 1550
QY 129 GERPRPDITKTVIVDDGLATGASMPAAVQALRDAPQAVIIVAPAA---PESTCREFA 185
Db 1551 EERQRT-PMRRYVLSLGSALLASSCM-----TVLAVVPARAQAQASTAMTTA 1596
QY 186 GLVDVVCATMPFPFLAVGESFMDPRQVTDDEVVRLLATPTAPSL----- 231
Db 1597 A---TSATPAARQIILLQARFVLQOOQYDNNARQALQNERIAPNRPDVLVGLGVOTAI 1652

QY 232 -RRPAATAADVLRRVAIDAPGVPT---HEVLAE-----LV 264
DB 1653 GNRBA---ADTLRLHQVAFCGSAAGNMLNDLSRAISQSDLSQIRSLAGSGQAQVA 1709
QY 265 GDARVLIGESSHC--THEFYQARAM--TOMLIEEGFCAVAADWDPAIRVN--RYVGL 321
DB 1710 GYOKLPHGKPKPHSLAVEYQTMAGVPAQWODARAGLGVVASN--PODYAQLAFAQAL 1767
QY 322 GEDTN-----ADEALSGERFPAMWMTYVVRDPE--WLRTT- 357
DB 1768 TYNSTRMEGLTRLKDLSPSQAPEVEAAAQSTRTQLSLMVPNPTQPLMEQHLNHP 1827
QY 358 -----NORYESGALRQAGFYGLD---LYSLHRSIOEVIY-----L 390
DB 1828 NDTLALREHMLHPGSPRDXAGLARQAGYQQLNAGSLAAEQSPQALQINSHDLSLGM 1887
QY 391 DKVDPRAAARARARYACDHAACADGQAYG--FAAFAQAGPSCEBAVEQLV-----DV 442
DB 1888 GLVSMRQGDTEAR--RYFEEBMAADPKTADRWPRPALAGAAVSGEYASVROLIAHQYTEA 1946
QY 443 ORNLAAVAROBGLADELFYA--QONQTVRDAEVVYRAMPSGRTVSMNLBDQMAQTL 500
DB 1947 KQOATTLARQPGQYTGATMLADLQRTGQIAAEQRYGLIS-----REPNNQAL 1998
QY 501 GSLTLHLDRHLDAPPARIVVAHNSHVGDARATEVWADGQTLGQIVRERYGDESRS--IG 559
DB 1999 MGL-----ARVDM-----ACQNTAEARQLLSRVPQYASQVG 2030
QY 560 FSTYGTIVTAASEWGIAQKRAVRPALHGSVEELFHQTADSFLVSARLSDAEAPLDVVR 619
DB 2031 EIEVSGMAAAQSQSDSARKVSI-----LREMAQAAPRDPBWVIRINLANLQCGQDVAE 2083
QY 620 LGRALGVVY---LPATERQSHYLVHRPADQPDAMHIDOTRALRELTSTWINGENET 676
DB 2084 AGRWOPILANPVTAQDRQAGILTYTSGSN--DAM-----TROLAGLSPAD 2128
QY 677 Y 677
DB 2129 Y 2129

RESULT 9
US-09-679-279-14
; Sequence 14, Application US/09679279
; Patent No. 6524841
; GENERAL INFORMATION:
; APPLICANT: McDaniel, Robert
; APPLICANT: Volchegursky, Yanina
; TITLE OF INVENTION: Recombinant Megalomicin Biosynthetic
; TITLE OF INVENTION: Genes and Uses Thereof
; FILE REFERENCE: 300622004700
; CURRENT APPLICATION NUMBER: US/09/679, 279
; CURRENT FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/158,305
; PRIOR FILING DATE: 1999-10-08
; PRIOR APPLICATION NUMBER: US 60/190,024
; PRIOR FILING DATE: 2000-03-17
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 3562
; TYPE: PRT
; ORGANISM: Micromonospora megalomicea
US-09-679-279-14

Query Match 3.8%; Score 134; Length 3562;
Best Local Similarity 24.1%; Pred. No. 0.0034;
Matches 172; Conservative 64; Mismatches 273; Indels 204; Gaps 40;

QY 47 ARGGIPVAMEVAALAHAPLDAFVVRKLGAP--GHDFAVGALASGGRVVVND--VVRGLR 103
DB 2580 AFGKLPFAMR--GVTLHA--TGATAVAVVATPAGPDAVALRVTDPTGQLVATVDALVVRDAG 2637

QY 104 ITPOQLDIAREGRELRLRESAYRGERPTDITGKTVIVVDGLATGASWFAAVALRD 163
DB 2638 ADROQPRG---RDS--DLHRLFWVLTATPPDT--PAAVVHAADGLD-----DLLRA 2682
QY 164 AQPQIVIAV---PAEESTCRFAGLV-----DDVCAQMPTEPLAVGESFWD 209
DB 2683 GGPAPQAVVVRVYRDGDDPTAEARHGVLMAATLVRRWLDD---RWPATTLVATVS--- 2735
QY 210 FROYDDEVRLTLATPTAGSLRPPASTADVLT-----RRVAID-----APGVPT 256
DB 2736 ---AGVEV-----SPGDDVFRPGAAVWGVLRCQAQSPRFLVVDGDPETPRPAVPD 2784
QY 257 HEVLAEVLGDAIV-----LIGESSHGTHEFYQARAAMTOMLIEEGFGAV--AAEADW 308
DB 2785 NPQAL--VRDQAVVPRLTPLAGVPVAVDRAVY-----LVPNG--SIEAVAPAV 2833
QY 309 PDAYR-----VNRVYVGLGEDTADALGSEFERFPAMWMTYVVRDPEVEMLRNRQYE 362
DB 2834 PDADRPLAPEEVRVAVRATG--VNFQVLLALGMYRPAEWGTEASGVTVESGVARFT 2891
QY 363 SGALRQ-----GFYGLDLYSLHRSIOEVIYSDKVDPPRA--ARARARYACFDHACAD 414
DB 2892 PG---QAVTGLFQCAFEPVAVADRLTLTPVDPGRVAVDAAVPLAFTTAHYALHDLAQLQ 2948
QY 415 DGQA--YGFPAAFGAGPS-----CEREAVE-----QLVDQORNLAVARODG-- 454
DB 2949 AGQSVLVHAAAGGVMAAVLARRAGAEVFATASPAKPTLRALGLDDHIASRESGFG 3008
QY 455 -----LAEDELFFYAQONQTVRDAEYTYAMSGR 485
DB 3009 ERFAARTGCGVDVVLNSLTGDLIDESARLLADGVE--VEGKTDLRPAE-----QFRGR 3062
QY 486 VTSNMLDQMAQVLTSLTH-----LDR-----HLDAAPPARIVVAHNSHVDA 530
DB 3063 YVPFDLAEAG--PDLGLSILSEVGLAAGALDRLPVSWELSAAPALYTHMSGRHVGKL 3121
QY 531 RATE---VMADGQL-----TLGQIVRERYGDESRSIGSTYGTVTAASEWG---GI 576
DB 3122 VLTQPAVHPDGTVLVVGSTIGRLV-----ARHLVYGHGPHLLVASRRPAPAGA 3174
QY 577 AQRRAVRPALHGSVEELFHQTADSFLVSARL--SRDAEAPLDVVRLGRALGVY 628
DB 3175 AELRADVEGIGATIEIVACDTRADREALALDISIPADRPL-----TGVVH 3219

RESULT 10
US-09-252-991A-17140
; Sequence 17140, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252, 991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 17140
; LENGTH: 535
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17140

Query Match 3.7%; Score 131; DB 2; Length 535;
Best Local Similarity 25.4%; Pred. No. 0.00028;
Matches 137; Conservative 44; Mismatches 193; Indels 166; Gaps 29;

QY 3 MTAADVTRSPRRVFRDREAGRVLAEILAAVYRDQVIVLGLARGLPVAME--VAAAL 61

Db LPAAGRARRRRPRVAPLRRLA--ARRLRIRPLPRRLAALLRAGAGGRMGLAAMP 93
Qy HAPL-----DAFVVRKLGAPGHDFAVAGALASGGRVVNDVVGRLITPQOLRDIARE 116
Db HVALAGSAAALERRRL-----RLAVGR-----ERRRGLR--PLRGHAPR 135
Qy GRELLRESAYRGERPPTDTGTGTIVVDDGLATGASMFPAVALRDAQIYI----- 171
Db GR--ARR--VQPRRLPADRVGRL-----GQPGANDLVGGRSLPGADPARPALRGKR 185
Qy AVPAABESTCRERAGLVDDVVCATMP-----TPFLAVGESFMPFQVTEDEVRRLATP 225
Db GTGVAPAAARRARA-----DV--ASLPRRRPRVPL-----RQLRRRPLP 225
Qy TAGPSLRPAASTADVLRRVAID-----APGVPTHEVLA 261
Db LRQDGLRAGPGLAESSRRRAAGRLTLAAGRTAAALLGAGVSLBSVDDHVALG 285
Qy E-----LVGDARIVLIGESSHGTHEFYQAPAAAT-----QMLIEKGFAY-- 302
Db HAAAGEDVARFDLIVGDPVAVLHLHLAGLH--YLAGAADHPAGLRVQAELEGGGVHG 343
Qy --AAEAD--WPAAYV--NRVYRGLEDNADDEALSGFERPPAMMRTVTVRDEVEM 353
Db FPGADADALAGGGEFHLVLDVDRYLRLG-----GF----- 374
Qy LRTNORYESGA-----LRQAGFYGLDLSLHRSIOEYISYDKVDPRAAR--ADARY 405
Db ---RGQPPVDGALDRRRRPGFVLVAAHVAVALQVVL-HRAHGRALAAEDPFLVEVRG 430
Qy ACPDHACADGQAY--GFAAFAFGAPSCEREAVQLVDVORNALA--YARQDGLAEDL 461
Db DQFQHAQVAVATVAGPFAAVALADDVQAEEVVDPRGHVQLAVVDDVLRRTGAVDEDHV 490

RESULT 11
US-08-426-630-49
Sequence 49, Application US/08426630
Patent No. 6656709
GENERAL INFORMATION:
APPLICANT: BLANCHE, FRANCIS; CAMERON, BEATRICE; CROUZET,
APPLICANT: JOEL; DEBUSCHE, LAURENT; LEVY SCHIL, SOPHIE;
APPLICANT: THIBAUT, DENIS
TITLE OF INVENTION: POLYPEPTIDES INVOLVED IN THE
TITLE OF INVENTION: BIOSYNTHESIS OF COBALAMINS AND/OR COBALAMIDES, DNA SEQUENCES
TITLE OF INVENTION: CODING FOR THESE POLYPEPTIDES, PREPARATION METHOD AND THEIR
TITLE OF INVENTION: USE.
NUMBER OF SEQUENCES: 60
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN
STREET: 555 13TH STREET, N.W.
CITY: WASHINGTON
STATE: DISTRICT OF COLUMBIA
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/426,630
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/916,151
FILING DATE: 14-SEP-1992
APPLICATION NUMBER: PCT/FR91/00054
FILING DATE: 30-JAN-1991
ATTORNEY/AGENT INFORMATION:
NAME: F. F. CALVETTI
REGISTRATION NUMBER: 28,557

REFERENCE/DOCKET NUMBER: 1290-7213
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 857-7887
TELEFAX: (202) 857-7929
INFORMATION FOR SEQ ID NO: 49:
SEQUENCE CHARACTERISTICS:
LENGTH: 1275 amino acids
TYPE: Amino acid
STRANDEDNESS:
TOPOLOGY: Linear
MOLECULE TYPE: protein
HYPOTHEICAL: No
ORIGINAL SOURCE:
ORGANISM: Pseudomonas denitrificans
STRAIN:
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPLOTYPE:
TISSUE TYPE:
CELL TYPE:
CELL LINE:
ORGANELLE:
FEATURE:
NAME/KEY: COBN
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION: Translation product of SEQ ID NO:48
US-08-426-630-49
Query Match 3.7%; Score 131; DB 2; Length 1275;
Best Local Similarity 20.0%; Pred. No. 0.0012;
Matches 127; Conservative 87; Mismatches 235; Indels 186; Gaps 29;
Qy LDAFVVRKLGAPGHDFAVAGALASGGRVVNDVVGRLITPQOLRDIAREBGRRLRE 124
Db LDAVLDLKEQMRDGLHIFGVAPBGRLLT--DITVALARVPRGLGEGGDSIQRAIAD 740
Qy 125 SAYRGERPPTDTGTGTIVVDDGLATGASMFPAVALRDAQP-----AQ 168
Db 741 AGRGFAIFTSAGNPA-----RDAQPDPLDCVMSDTWTGPKPS 780
Qy 169 IYAVAAPESTCRERAGLVDDVVCATMPFPLAVGESFMPFQVTEDEVRRLATPTAG 228
Db 781 ILADLDAPWRTGIDVVERI-ELLANLVSGELACPDHMANTAVALGEIETRL-----K 833
Qy 229 PSLRPAASTAADVLR-----RVAIDAPGVPTHEVLAELVGDARIVLIGESSHGTHEY- 283
Db 834 PSISNGALAMTGLTGLSGRFVAPGSCAPTR-----GRPDVLPTRGN-----FIS 880
Qy 284 -QARAAMTQWLEBKGFVAVAE-----ADWPDAYRVNRV-----VRGLGEDTNAD 328
Db 881 VDSRAVPTPAAYE--LGKSAELLIRRYLDHGEWPSFGLTAWGTAMRTGGDDIAQA 937
Qy 329 EALSGERPPAMMRTVTVRDFVEMLTRNQRRE-----SGALRQAGFGDL 376
Db 938 LALIGAK--FTW-----DWSRRVMGYEIVPLAVLGRPRVDVLTLSGFF-RDA 983
Qy 377 YS-----LHRSIOEYISYDKVDPRAAARARAYACFDHACDDGAYGPA--AFGAGP 429
Db 984 FPDQIALFDPAIVAVLAEEDDADNMIAARRAESRLEBGEAAEARRASTRVGAKP 1043
Qy 430 SCERAEVQLVD-----VORNALAYARQDGLAEDLIFYAQNAQTVR--- 472
Db 1044 GAYGALLQALIDKEKMETKADLAELVTCGAYVYVAGBEGSKARDLF--EERLRTIYAV 1101
Qy 473 -----DAEVYIRAMFSGRVTSWNLRDQMAQTGLSLTHLDRLHDAIPARIYVW 521
Db 1102 QONDNREHLLDSDDYQ--FEGGMSA-----AAEOLGHRPALYHNDHSRPERPVIR 1152
Qy 522 AHSNHYGD-ARATEV--WADQGLTIGQIVREHYGDSHSIGFSTYGTVAASE-----W 573
Db 1153 SLEEEIGRVRAVAVPKWIDGV-----RHG-----YKGAPEIATVADYMF 1194

Qy 574 GGIQRKAVRPAALHSGVSEELFHQTADSFLVSARLS 608
Db 1195 AFAATTGAVR-----DHHEFAAYQAFIVDERVA 1222

RESULT 12

US-09-410-551B-72
; Sequence 72, Application US/09410551B
; Patent No. 6503737
; GENERAL INFORMATION:
; APPLICANT: KOSAN BIOSCIENCES, Inc.
; APPLICANT: REEVES, CHRISTOPHER
; APPLICANT: CHU, DANIEL
; APPLICANT: KHOSLA, CHAITAN
; APPLICANT: SANTI, DANIEL
; APPLICANT: WU, KAI
; TITLE OF INVENTION: POLYKETIDE SYNTHASE ENZYMES AND RECOMBINANT DNA
; TITLE OF INVENTION: CONSTRUCTS THEREFOR
; FILE REFERENCE: 30062-20026.00
; CURRENT APPLICATION NUMBER: US/09/410,551B
; PRIOR FILING DATE: 1999-10-01
; PRIOR APPLICATION NUMBER: US 60/139,650
; PRIOR FILING DATE: 1999-06-17
; PRIOR APPLICATION NUMBER: US 60/123,810
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 60/102,748
; PRIOR FILING DATE: 1998-10-02
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PASTESEQ for Windows Version 4.0
; SEQ ID NO 72
; LENGTH: 6396
; TYPE: PRT
; ORGANISM: Streptomyces hygroscopicus
; US-09-410-551B-72

Query Match 3.7%; Score 130; DB 2; Length 6396;
Best Local Similarity 21.0%; Pred. No. 0.024;
Matches 163; Conservative 64; Mismatches 274; Indels 274; Gaps 31;

Qy 17 VFRRREAGR-----VLAELAAAYRD-----QPDVIVGLARGG----- 50
Db 4453 MFRGMRAMWDGDTVAEVALPEDRAADADFGGHPALDAAALQSSGLMLLESDEGSVQ 4512
Qy 51 LPVAWEVAALAHAPLDFAVVRKLGAPGHDEFAGVALASGGRVVVNDVVRGLRTTPOOLR 110
Db 4513 LPFSWM- GVRFA- TGATMLRVAVVRGPDGLRLHAADSGNRPAVTIDAL-----VTRSPEA 4566
Qy 111 DIAR-----EGRELLRR--ESAYRGERP 132
Db 4567 DLAPADMLRVGMAPVVPVAGAGPSDADVLTLRGDDADPLGETRDLTTRVLDALLRADRP 4626
Qy 133 PTDTGKTIVVDDGLATGASMPFAVQALRDAPQAVIAVPAAPESICREFAGLVDDV 192
Db 4627 -----VIRQVVTGLAAKAAA--GLVRTAQNEQPGRRFL--VETDP-----GEVL 4666
Qy 193 CATMPTPELAVGSEFMDFRQVTDSEVRRLATPT-----AGSLRRPAASTAADVLRRVA 247
Db 4667 DGAKRDALALGEPHYVLRDLFEARILMRATPSLTLPDGTGSKQLRBSATGSLDDLAIVP 4726
Qy 248 IDAB-----GGVPTHEVLAEL--VGDA-----IVLIGESSHGTHF 282
Db 4727 TDADDRPLAGEVRIAVRAAGLNFPRDVTVALGVVADARPLGSBAAGVVL--ETGPGVHDL 4784
Qy 283 YQAPAAATQMTLIERKGCAGAVAAEADWPAYRVNRYVKGLEDYTADELTSFERF-PAWM 341
Db 4785 APGRVLT-----GMLAGAFGPVALITDRRLIGMPGKWTFPQAAASVMTAFATAM- 4832
Qy 342 WRNTVVDPEVEMLTRNQRYESGALRQAGFYGL-DIVSIHRSIOEVISYLDKVDPRRAAR 400
Db 4833 -----VIRQVVTGLAAKAAA--GLVRTAQNEQPGRRFL--VETDP-----GEVL 4666
Qy 401 ARAR-----YAC-----FDHACADDGQAYGPAAR----- 425

Db 4862 QIANHLGAERYATTSAAKRHLVDLDGNAHLDNSTASTADAFPPVDVVLNSITGEFLDASV 4921
Qy 426 -----GAGPSCREAVEQLVDV-----QRNALAYARODGLAEDELFYAQONAOQT 470
Db 4922 GLLAAGGRPIEMKGTDIRHAVQGPFDLMDAGPRMQRILYBLGLFARVDVLHPLFPVAMWD 4981
Qy 471 VRDAEVYTRAMFSGRVTSMNLRODHMAQTGLSLTTHLDRLHDAAPARIYVMAHNSHVDA 530
Db 4982 VRCAREAFGMWSSGRHT-----GKLVLPFRPLD--PEGAVVI----- 5017
Qy 531 RATEVMADGQLTQIYRERYGDE-----SRSIGFTYGT----- 566
Db 5018 -----TGSGGTLAGILARHLGHPHYLLSRTPPTTGGTHLPCEVDGPHQATTLARI 5071
Qy 567 ---VTAASEWGIAQRKAVRPAALHSGVSEELFHQTADSFLVSARLSRDAEALDVV 618
Db 5072 PQPLTAVFHTAGTIDDLNLNLPDRVDVYLKPRADAAHMLHRLTRDTDLAAFAVV 5126

RESULT 13
US-09-940-316B-72
; Sequence 72, Application US/09940316B
; Patent No. 6759536
; GENERAL INFORMATION:
; APPLICANT: KOSAN BIOSCIENCES, Inc.
; APPLICANT: REEVES, CHRISTOPHER
; APPLICANT: CHU, DANIEL
; APPLICANT: KHOSLA, CHAITAN
; APPLICANT: SANTI, DANIEL
; APPLICANT: WU, KAI
; TITLE OF INVENTION: POLYKETIDES ENCODING THE FKBA GENE OF THE FK-520 POLYKETIDE SYNTHASE
; TITLE OF INVENTION: GENE CLUSTER
; FILE REFERENCE: 30062-20026.11
; CURRENT APPLICATION NUMBER: US/09/940,316B
; CURRENT FILING DATE: 2001-08-27
; PRIOR FILING DATE: 1999-10-01
; PRIOR APPLICATION NUMBER: US 60/139,650
; PRIOR FILING DATE: 1999-06-17
; PRIOR APPLICATION NUMBER: US 60/123,810
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 60/102,748
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PASTESEQ for Windows Version 4.0
; SEQ ID NO 72
; LENGTH: 6396
; TYPE: PRT
; ORGANISM: Streptomyces hygroscopicus
; US-09-940-316B-72

Query Match 3.7%; Score 130; DB 2; Length 6396;
Best Local Similarity 21.0%; Pred. No. 0.024;
Matches 163; Conservative 64; Mismatches 274; Indels 274; Gaps 31;

Qy 17 VFRRREAGR-----VLAELAAAYRD-----QPDVIVGLARGG----- 50
Db 4453 MFRGMRAMWDGDTVAEVALPEDRAADADFGGHPALDAAALQSSGLMLLESDEGSVQ 4512
Qy 51 LPVAWEVAALAHAPLDFAVVRKLGAPGHDEFAGVALASGGRVVVNDVVRGLRTTPOOLR 110
Db 4513 LPFSWM- GVRFA- TGATMLRVAVVRGPDGLRLHAADSGNRPAVTIDAL-----VTRSPEA 4566
Qy 111 DIAR-----EGRELLRR--ESAYRGERP 132
Db 4567 DLAPADMLRVGMAPVVPVAGAGPSDADVLTLRGDDADPLGETRDLTTRVLDALLRADRP 4626
Qy 133 PTDTGKTIVVDDGLATGASMPFAVQALRDAPQAVIAVPAAPESICREFAGLVDDV 192
Db 4627 -----VIRQVVTGLAAKAAA--GLVRTAQNEQPGRRFL--VETDP-----GEVL 4666
Qy 193 CATMPTPELAVGSEFMDFRQVTDSEVRRLATPT-----AGSLRRPAASTAADVLRRVA 247

Db 4667 DGAKRDALALGEBHVRLROGLFEARILMRATPSLTLPDTGSKQLRPSATGSLDDLAIVP 4726
Qy 248 IDAP-----GGVPTHEVLAEL--VGDAE-----IVLIGESSHGTHF 282
Db 4727 TDADPRPLAGEVRLAIVRAAGLNFEDVTVALGVADAPPLGSEAAVGL--ETGPGVYDL 4784
Qy 283 YQAAQAATQMLIEERKGAVALAEDWPRAYRVNRVYRGSGEDTADELSGFEF--PAMM 341
Db 4785 APGRVL-----GMLAGAFGPVATIDRRLGMPDGWTFPPQAAVMTAFATAM- 4832
Qy 342 WRNTVVDVFEVMTLRNQRYESGALRQAGFYGL-DLYSLHRSIOBVISYLDKVPRAAR 400
Db 4833 -----YGLVDLAGL-RPGKVLHHAATGVGAAY 4861
Qy 401 ARAR-----YAC-----FDHACADDGAYGFAAF-----425
Db 4862 QIARHLGAEVYATTSAAKRLVLDGAILDSRSTAFADAPPVVDVNLSTGFEFLDASV 4921
Qy 426 -----GAGSCREAVEQLVDV-----QRNALAYARQDGLADELFYAQONAQ 470
Db 4922 GLLAAGRFLEMGTDIRHAHQDPDLMADAGPDMQRIYELLGLFADVHLPLPVHAMW 4981
Qy 471 VRDAEYVYRAMFGSRTVSMNLRDQMAQTGLSLTLTHDRHLADAPPARIVWAHNSHYGDA 530
Db 4982 VRQAREAFGMSGRHT-----SKLVLTVPRLD-PEGAVI-----5017
Qy 531 RATEWMDGOLLGQIVRERYGDE-----SRISGSTYTG-----566
Db 5018 -----TGSGGTLAGILARHLGHPHTYLSRTPPDTPGTHLPCVDGPHQLATTLARI 5071
Qy 567 ---VTAASBNGIQRKAVNRPALHGSVELFHQTADSEVLSARLSRAEAPLDVY 618
Db 5072 POPLAVFHTAGTLDLALDNLTPRDVTVLKPRADAMHLRLTRTDLAFAVV 5126

RESULT 14
US-09-477-962-117
; Sequence 117, Application US/09477962
; Patent No. 6927286
; GENERAL INFORMATION:
; APPLICANT: SHEN, BEN
; APPLICANT: DU, LIANGCHENG
; APPLICANT: SANCHEZ, CESAR
; APPLICANT: CHEN, MEI
; APPLICANT: EDWARDS, DANIEL J.
; TITLE OF INVENTION: BLEOMYCIN GENE CLUSTER COMPONENTS AND THEIR USES
; FILE REFERENCE: 407T-895820US
; CURRENT FILING DATE: 2000-01-05
; PRIOR APPLICATION NUMBER: US/09/477,962
; PRIOR FILING DATE: 1999-01-06
; PRIOR APPLICATION NUMBER: 60/115,435
; PRIOR FILING DATE: 1999-02-05
; NUMBER OF SEQ ID NOS: 133
; SOFTWARE: Patentin Ver. 3.0
; SEQ ID NO 117
; LENGTH: 2841
; TYPE: PRT
; ORGANISM: Streptomyces verticillius
; FEATURE:
; OTHER INFORMATION: ORF32
US-09-477-962-117

Query Match 3.7%; Score 128.5; DB 2; Length 2841;
Best Local Similarity 20.9%; Pred. No. 0.0086;
Matches 158; Conservative 70; Mismatches 263; Indels 265; Gaps 36;

Qy 4 TAAADVTTR-----SPRVFRDRREAGRVLAE-LTAAVVDQPDVTVLGIANGGLP-----52
Db 1966 TAAAPVITRLRLPADPGVHLDRQADRDRLRLHRDQAAYITFTSGSGAKGKVVR 2025
Qy 53 -----VAMEVAAALHAPLDAFVVRKUGAPGH-----DEFA 82

Db 2026 HGSLYHLGHVRMAEGCPRRNVANTTAMTFDPSLEFMLVAGHTLHAPEVRDPEA 2085
Qy 83 VGALASGRRVNDVVRGRLRTTPQOLRIABRBGELLRESAYNGERPPTDTIGKTVI 142
Db 2086 LVALLVRRAIDV-----LNTVPSHL-----TLTIEAGLIEGDRVP-----GTVL 2124
Qy 143 VVDDGL-----ATGA-----SMPAAVQALRDA-----QPAQVIAVPAAPESTCR 182
Db 2125 VGBEAVPAALMRLTRERTGTREFNLVGPTEAVTADCHDLSDPADVPVGTPLPHTHVR 2184
Qy 183 EFAGLVDDVVCATMPFPFLVGSF-----207
Db 2185 -----VLD-----RLRPVPGVAGELIYLGCTGLARGYLNRPALTQORFADPYPTPGSRLY 2237
Qy 208 -----W-----DFQVTDSEVR-----RLIATPTAGPSLRRAASTAAD--VL 243
Db 2238 RTGDRAMWRBDGTLEYLGRTRDDQIKIRGFVPEGEIEAVLTHHRAVEAAVVAATDGA 2297
Qy 244 RRVALDAPGVPTHEVLAELVGDAKIVLIGESSHGHEFYQAAAMTQMLIEERGF-----239
Db 2298 RLVALVPAAPAPHGSADGAPDAQV-----EEMNAVFEATHTDAAGELTFENIKGWDLSL 2353
Qy 300 -GA-VAAE--ADWP-----AYRVNRVYRGSGEDTNADEALSGFERPAMMR--343
Db 2354 TGAPIRAEHRKENVDTTVALRLERPARVLEIGSGTG-----LIMKRL 2397
Qy 344 -----NTVVDVFEVMTLRNQRYESGALR-----QAGFYGL-----DLYSLHRSIOB 385
Db 2398 PHTEYTGTFSPRAVDWMLDGLRRRPARHRVRLHREATDFTGVRAASTDLVVVNSVQY 2457
Qy 386 V--ISYLDKVPRAARARARVACFDHACDDG-----QAYGFAAAGAP 429
Db 2458 FPDRAVLDTVLARAL-----DATADRGRVFGVDVRNIALAPQFAAQALAHAGP 2506
Qy 430 SCREAVEQLVDVQRNALAYARQDGLADELFYAQONAQTVDAEYVYRAMFGSRTSW 489
Db 2507 GAAR-----DARAAGEPAAAMDGLVSPAYFALAAAR-----SPRTGV 2547
Qy 490 NL--BDQH--MAQTGLSLTLH--DRHLADAPPARIVV--AHNSHVGADARATEVMAD 538
Db 2548 EILPRGRHNRNEMSLRVYDVVLHVGGDRPA-ADEAEVLGTGDDQVHLASISARLGGGPD 2606
Qy 539 GQLTGQIVRERYGDSRSIGFSTYGTVTAASENG 574
Db 2607 ALLVRG-VANDRLTRDNEILDAPARTTAVPEPDLWG 2641

RESULT 15
US-10-042-665A-8
; Sequence 8, Application US/10042665A
; Patent No. 6924106
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Toupet, Christine
; APPLICANT: Engel, Nathalie
; TITLE OF INVENTION: Rifamycin biosynthesis gene cluster
; FILE REFERENCE: 4-21001/B/C1
; CURRENT FILING DATE: 2002-01-09
; PRIOR APPLICATION NUMBER: US/10/042,665A
; PRIOR FILING DATE: 1999-03-24
; PRIOR APPLICATION NUMBER: PCT/EP97/04495
; PRIOR FILING DATE: 1997-08-18
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 8
; LENGTH: 3413
; TYPE: PRT
; ORGANISM: Amycolatopsis mediterranei
US-10-042-665A-8

Query Match 3.7%; Score 128.5; DB 2; Length 3413;
Best Local Similarity 21.3%; Pred. No. 0.012;

THIS PAGE BLANK (USPTO)

This Page Blank (USPTO)


```
Dh 301 AWAADWPAAYRVNRYRGIGEDTNADEALSGFERPPAMMNTVVRDVEMLTRNQR 360
Qy 361 YESGALROAGFYGLDLYSLHRSIQEVI SYLDKVPRAAARARAYACEPDHACADGQAYG 420
Dh 361 YESGALROAGFYGLDLYSLHRSIQEVI SYLDKVPRAAARARAYACEPDHACADGQAYG 420
Qy 421 FAAAFGAGPSCERAEVQVLDVQORNAALAYARQDGLAEDELFTYQONNQTVDRAEVYTRA 480
Dh 421 FAAAFGAGPSCERAEVQVLDVQORNAALAYARQDGLAEDELFTYQONNQTVDRAEVYTRA 480
Qy 481 MFSGRVTSWNLRDQHMOTGSLTLHLDRHLDAAPARIVVAHNSHGDARATEVMADGQ 540
Dh 481 MFSGRVTSWNLRDQHMOTGSLTLHLDRHLDAAPARIVVAHNSHGDARATEVMADGQ 540
Qy 541 LTTGQIVRERYGDSRSIGFSTYGTVVAASEMGIAQORKAVPALHSGVELFHQTADS 600
Dh 541 LTTGQIVRERYGDSRSIGFSTYGTVVAASEMGIAQORKAVPALHSGVELFHQTADS 600
Qy 601 FLVSARLSRDAEAPLDVVRGRLAIGVYVLPATERQSHYLVPRADQPDAMTHIDQTRALE 660
Dh 601 FLVSARLSRDAEAPLDVVRGRLAIGVYVLPATERQSHYLVPRADQPDAMTHIDQTRALE 660
Qy 661 PLEVTSRWIAGENETYPPTGL 681
Dh 661 PLEVTSRWIAGENETYPPTGL 681
```

RESULT 2

```
US-11-045-942-10
; Sequence 10, Application US/11045942
; Publication No. US20050172350A1
; GENERAL INFORMATION:
; APPLICANT: Williams, Deryck Jeremy
; APPLICANT: Salmon, Brian
; TITLE OF INVENTION: NEMATODE PPPT-LIKE SEQUENCES
; FILE REFERENCE: 12557-005001
; CURRENT APPLICATION NUMBER: US/11/045,942
; PRIOR FILING DATE: 2005-01-27
; PRIOR APPLICATION NUMBER: US/10/113,201
; PRIOR FILING DATE: 2002-03-29
; PRIOR APPLICATION NUMBER: US 60/280,192
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 443
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis PPPT
US-11-045-942-10
```

```
Query Match 12.3%; Score 431; DB 6; Length 443;
Best Local Similarity 32.4%; Pred. No. 1.8e-29;
Matches 129; Conservative 52; Mismatches 161; Indels 56; Gaps 11;
```

```
Qy 16 RVFRDRREAGRVLAELLAAYRDQPDVIVGLARGGLPVAMEVAAALHAPLDAFVVRKLG 75
Dh 2 KLFDKRGDAQRQLAQLAQLSGKA-VVVLGLPRGVPVAFVFAKSLQAPLDVLRKLG 60
Qy 76 PGHDEPVALGASGRVAVNDVVRGLRITPQQLRDLAEREGELLRESAYGERPPTD 135
Dh 61 PPOBELAFGAIGEDGVAVLNDVVRGTHLDAAMDVERKQILQLQRAERFRGRDRIP 120
Qy 136 ITGKTIVVDDGLATGASMPAAVOALRDAPQIVIVPAAPESTCEPAGLVDDVVCAT 195
Dh 121 LTGRIAVIVDDGIATGATATAACQVAAHAGADKVLAVIPGPDIVARFAGYADDEVCLA 180
Qy 196 MPTPFLAVGESFMDPQVOTDEEVRRLLATPTAGPSLRPAASTADVLRVAIDAPGVP 255
Dh 181 TPALFFAVGGGYRFTQTSDEVAFL-----DRAHRDFAEKAGATDAADDP 227
Qy 256 THEVLAEVGDARIVLIGESSHGTHERYQARAAMTOMLIEKFGAVA-AEADWPDAYRV 314
Dh 325 TASLPVGYFG-----ASTGAGAAALVAAADPFRVNVRA 355
```

```
Dh 228 LRDEEVUVAGP-----VPVAGHLT-----VEKPRGIVVFAHSGSSRHSI 269
Qy 315 -NRVY-----RGLGEBTNADALSGFER-----FPAMMNTVVRDVEMLTRNQRYES 363
Dh 270 RNRVYAEVLVAGAPATLLFOLLTPBEERBNANVFDEILASRLI-DVTGWLATQ----PD 324
Qy 364 GALRQAGFYGLDLYSLHRSIQEVI SYLDKVPRAAARA 401
Dh 325 TASLPVGYFG-----ASTGAGAAALVAAADPFRVNVRA 355
```

RESULT 3

```
US-11-058-869-10
; Sequence 10, Application US/11058869
; Publication No. US20050172363A1
; GENERAL INFORMATION:
; APPLICANT: Kloeck, Andrew P.
; APPLICANT: Williams, Deryck Jeremy
; TITLE OF INVENTION: NEMATODE PPPT-LIKE SEQUENCES
; FILE REFERENCE: 12557-005001
; CURRENT APPLICATION NUMBER: US/11/058,869
; PRIOR FILING DATE: 2005-02-16
; PRIOR APPLICATION NUMBER: US/10/113,201
; PRIOR FILING DATE: 2002-03-29
; PRIOR APPLICATION NUMBER: US 60/280,192
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 443
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis PPPT
US-11-058-869-10
```

```
Query Match 12.3%; Score 431; DB 6; Length 443;
Best Local Similarity 32.4%; Pred. No. 1.8e-29;
Matches 129; Conservative 52; Mismatches 161; Indels 56; Gaps 11;
```

```
Qy 16 RVFRDRREAGRVLAELLAAYRDQPDVIVGLARGGLPVAMEVAAALHAPLDAFVVRKLG 75
Dh 2 KLFDKRGDAQRQLAQLAQLSGKA-VVVLGLPRGVPVAFVFAKSLQAPLDVLRKLG 60
Qy 76 PGHDEPVALGASGRVAVNDVVRGLRITPQQLRDLAEREGELLRESAYGERPPTD 135
Dh 61 PPOBELAFGAIGEDGVAVLNDVVRGTHLDAAMDVERKQILQLQRAERFRGRDRIP 120
Qy 136 ITGKTIVVDDGLATGASMPAAVOALRDAPQIVIVPAAPESTCEPAGLVDDVVCAT 195
Dh 121 LTGRIAVIVDDGIATGATATAACQVAAHAGADKVLAVIPGPDIVARFAGYADDEVCLA 180
Qy 196 MPTPFLAVGESFMDPQVOTDEEVRRLLATPTAGPSLRPAASTADVLRVAIDAPGVP 255
Dh 181 TPALFFAVGGGYRFTQTSDEVAFL-----DRAHRDFAEKAGATDAADDP 227
Qy 256 THEVLAEVGDARIVLIGESSHGTHERYQARAAMTOMLIEKFGAVA-AEADWPDAYRV 314
Dh 228 LRDEEVUVAGP-----VPVAGHLT-----VEKPRGIVVFAHSGSSRHSI 269
Qy 315 -NRVY-----RGLGEBTNADALSGFER-----FPAMMNTVVRDVEMLTRNQRYES 363
Dh 270 RNRVYAEVLVAGAPATLLFOLLTPBEERBNANVFDEILASRLI-DVTGWLATQ----PD 324
Qy 364 GALRQAGFYGLDLYSLHRSIQEVI SYLDKVPRAAARA 401
Dh 325 TASLPVGYFG-----ASTGAGAAALVAAADPFRVNVRA 355
```

RESULT 4

```
US-11-110-578-10
; Sequence 10, Application US/11110578
; Publication No. US20050186631A1
; GENERAL INFORMATION:
```



```

; ORGANISM: Meloidogyne incognita PPPT
US-11-045-942-4

Query Match      11.9%; Score 417.5; DB 6; Length 233;
Best Local Similarity 41.9%; Pred. No. 1.1e-28;
Matches 88; Conservative 44; Mismatches 75; Indels 3; Gaps 2;

OY 12 RSPRRVRDRREARARVLAELLAAARQDPD-VIVYGLARGGLPVAMVEYAAALHAFLDAFV 70
DB 5 RAAATAPFDRDRDAQKLAELAKNFKSORDKVVVLTALPRGGVPVAFVAKSIGAPLDLIMV 64
OY 71 RKLGAPEHDEFAVAGALASGG--RVVNVDDVVRGLRITPOOLRDIAREGRELRLRESAYR 128
DB 65 RKIAGAPHEHEGIGAVVEGNPPELVNMEDAVKTYQPEBGVYQAMMEQKLEIARQRRL 124
OY 129 GERPDITIGKTVIVVDGIGATGASMEFAVQALRDAOPAQIVIAVPAPESTCEFFAGLV 188
DB 125 GDRPEPLSIAGKIAIVVDGIGATGTAIVAMKALRQKVAKALLASPLAPSDTLAELRAEG 184
OY 189 DDVVCATMPTPEFLAVGESFMDPQVNTDEEV 218
DB 185 NEVLVLETPEPNFSAVGLHYTKFDQTSDEEV 214

RESULT 6
US-11-058-869-4
; Sequence 4, Application US/11058869
; Publication No. US20050172363A1
; GENERAL INFORMATION:
; APPLICANT: Kloek, Andrew P.
; APPLICANT: Williams, Deryck Jeremy
; APPLICANT: Salmon, Brian
; TITLE OF INVENTION: NEMATOIDE PPPT-LIKE SEQUENCES
; FILE REFERENCE: 12557-005001
; CURRENT APPLICATION NUMBER: US/11/058, 869
; PRIOR FILING DATE: 2005-02-16
; PRIOR APPLICATION NUMBER: US/10/113, 201
; PRIOR FILING DATE: 2002-03-29
; PRIOR APPLICATION NUMBER: US 60/280, 192
; PRIOR FILING DATE: 2001-03-30
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 233
; TYPE: PRT
; ORGANISM: Meloidogyne incognita PPPT
US-11-058-869-4

Query Match      11.9%; Score 417.5; DB 6; Length 233;
Best Local Similarity 41.9%; Pred. No. 1.1e-28;
Matches 88; Conservative 44; Mismatches 75; Indels 3; Gaps 2;

OY 12 RSPRRVRDRREARARVLAELLAAARQDPD-VIVYGLARGGLPVAMVEYAAALHAFLDAFV 70
DB 5 RAAATAPFDRDRDAQKLAELAKNFKSORDKVVVLTALPRGGVPVAFVAKSIGAPLDLIMV 64
OY 71 RKLGAPEHDEFAVAGALASGG--RVVNVDDVVRGLRITPOOLRDIAREGRELRLRESAYR 128
DB 65 RKIAGAPHEHEGIGAVVEGNPPELVNMEDAVKTYQPEBGVYQAMMEQKLEIARQRRL 124
OY 129 GERPDITIGKTVIVVDGIGATGASMEFAVQALRDAOPAQIVIAVPAPESTCEFFAGLV 188
DB 125 GDRPEPLSIAGKIAIVVDGIGATGTAIVAMKALRQKVAKALLASPLAPSDTLAELRAEG 184
OY 189 DDVVCATMPTPEFLAVGESFMDPQVNTDEEV 218
DB 185 NEVLVLETPEPNFSAVGLHYTKFDQTSDEEV 214

RESULT 7
US-11-110-578-4
; Sequence 4, Application US/1110578
; Publication No. US20050186631A1

```


; PRIOR FILING DATE: 2001-03-30
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 229
; TYPE: PRT
; ORGANISM: Meloidogyme javanica PPPT
US-11-058-869-5

Query Match 11.4%; Score 401.5; DB 6; Length 229;
Best Local Similarity 40.2%; Pred. No. 3e-27;
Matches 84; Conservative 45; Mismatches 77; Indels 3; Gaps 2;

QY 17 VFDRRREAGRYLAELLAAYRDQPD-VIYGLARGGLPVAMEVAAALHAFLDAFVVRKIGA 75
D 9 LFIDRKDAGQKLAALAHILPQRDNIVVIALPRGGVPVACEADAFAQAPLDLMLVRKIGA 68
QY 76 PGHDEPAVAGALASG--RVVYVNDVVRGLRTTPOQLDIAREBERELLRRESAYRGERPP 133
D 69 PGHEEYIGAVVEGNPELVNNEBAVRITRPESEAYVQOQOKOLEMEROKRTYLGDKPP 128
QY 134 TDIGKTVIYVDDGLATGASMPAAVQALRDAQPAQIVIAVPAPESTCREFAGLVDDVVC 193
D 129 VSLGRITAIYVDDGIATGATGRVAKLKARQKNVSRAILASPMAPSDTLAEIRAGNEVLC 188
QY 194 ATMPTPLAVGSEFWDFRQVTDDEVRRL 222
D 189 LETPNNFSAVGLHYORFQTSDEVRICM 217

RESULT 11
US-11-058-869-6
; Sequence 6, Application US/11058869
; Publication No. US20050172363A1
; GENERAL INFORMATION:
; APPLICANT: Kloeck, Andrew P.
; APPLICANT: Williams, Deryck Jeremy
; TITLE OF INVENTION: NEMATODE PPPT-LIKE SEQUENCES
; FILE REFERENCE: 12557-005001
; CURRENT APPLICATION NUMBER: US/11/058,869
; CURRENT FILING DATE: 2005-02-16
; PRIOR APPLICATION NUMBER: US/10/113,201
; PRIOR FILING DATE: 2002-03-29
; PRIOR APPLICATION NUMBER: US 60/280,192
; PRIOR FILING DATE: 2001-03-30
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 229
; TYPE: PRT
; ORGANISM: Heterodera glycines PPPT
US-11-058-869-6

Query Match 11.4%; Score 401.5; DB 6; Length 229;
Best Local Similarity 40.2%; Pred. No. 3e-27;
Matches 84; Conservative 45; Mismatches 77; Indels 3; Gaps 2;

QY 17 VFDRRREAGRYLAELLAAYRDQPD-VIYGLARGGLPVAMEVAAALHAFLDAFVVRKIGA 75
D 9 LFIDRKDAGQKLAALAHILPQRDNIVVIALPRGGVPVACEADAFAQAPLDLMLVRKIGA 68
QY 76 PGHDEPAVAGALASG--RVVYVNDVVRGLRTTPOQLDIAREBERELLRRESAYRGERPP 133
D 69 PGHEEYIGAVVEGNPELVNNEBAVRITRPESEAYVQOQOKOLEMEROKRTYLGDKPP 128
QY 134 TDIGKTVIYVDDGLATGASMPAAVQALRDAQPAQIVIAVPAPESTCREFAGLVDDVVC 193
D 129 VSLGRITAIYVDDGIATGATGRVAKLKARQKNVSRAILASPMAPSDTLAEIRAGNEVLC 188
QY 194 ATMPTPLAVGSEFWDFRQVTDDEVRRL 222
D 189 LETPNNFSAVGLHYORFQTSDEVRICM 217

RESULT 12
US-11-110-578-5
; Sequence 5, Application US/11110578
; Publication No. US20050186631A1
; GENERAL INFORMATION:
; APPLICANT: Kloeck, Andrew P.
; APPLICANT: Williams, Deryck Jeremy
; TITLE OF INVENTION: NEMATODE PPPT-LIKE SEQUENCES
; FILE REFERENCE: 12557-005001
; CURRENT APPLICATION NUMBER: US/11/110,578
; CURRENT FILING DATE: 2005-04-20
; PRIOR APPLICATION NUMBER: US/10/113,201
; PRIOR FILING DATE: 2002-03-29
; PRIOR APPLICATION NUMBER: US 60/280,192
; PRIOR FILING DATE: 2001-03-30
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 229
; TYPE: PRT
; ORGANISM: Meloidogyme javanica PPPT
US-11-110-578-5

Query Match 11.4%; Score 401.5; DB 6; Length 229;
Best Local Similarity 40.2%; Pred. No. 3e-27;
Matches 84; Conservative 45; Mismatches 77; Indels 3; Gaps 2;

QY 17 VFDRRREAGRYLAELLAAYRDQPD-VIYGLARGGLPVAMEVAAALHAFLDAFVVRKIGA 75
D 9 LFIDRKDAGQKLAALAHILPQRDNIVVIALPRGGVPVACEADAFAQAPLDLMLVRKIGA 68
QY 76 PGHDEPAVAGALASG--RVVYVNDVVRGLRTTPOQLDIAREBERELLRRESAYRGERPP 133
D 69 PGHEEYIGAVVEGNPELVNNEBAVRITRPESEAYVQOQOKOLEMEROKRTYLGDKPP 128
QY 134 TDIGKTVIYVDDGLATGASMPAAVQALRDAQPAQIVIAVPAPESTCREFAGLVDDVVC 193
D 129 VSLGRITAIYVDDGIATGATGRVAKLKARQKNVSRAILASPMAPSDTLAEIRAGNEVLC 188
QY 194 ATMPTPLAVGSEFWDFRQVTDDEVRRL 222
D 189 LETPNNFSAVGLHYORFQTSDEVRICM 217

RESULT 13
US-11-110-578-6
; Sequence 6, Application US/11110578
; Publication No. US20050186631A1
; GENERAL INFORMATION:
; APPLICANT: Kloeck, Andrew P.
; APPLICANT: Williams, Deryck Jeremy
; TITLE OF INVENTION: NEMATODE PPPT-LIKE SEQUENCES
; FILE REFERENCE: 12557-005001
; CURRENT APPLICATION NUMBER: US/11/110,578
; CURRENT FILING DATE: 2005-04-20
; PRIOR APPLICATION NUMBER: US/10/113,201
; PRIOR FILING DATE: 2002-03-29
; PRIOR APPLICATION NUMBER: US 60/280,192
; PRIOR FILING DATE: 2001-03-30
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 229
; TYPE: PRT
; ORGANISM: Heterodera glycines PPPT
US-11-110-578-6

Query Match 11.4%; Score 401.5; DB 6; Length 229;
Best Local Similarity 40.2%; Pred. No. 3e-27;

Db 1257 GALPALLERGLVAPIRATGAGVEDESSLGRMAAMPVAREBQVLVDVTRQVATVIGH 1316
Qy 193 -CATMPTFFLAVGSEFWDFQVTDSEVR-----RLATP----- 225
Db 1317 GTATRAVDPAATFAETGFD--SLTAVELNRLRTATGVLSATAFDYPTPAVLGHLLRE 1374
Qy 226 ---TAGSELRPPAATAAD-----VLRRVADAPGCPTEHEVLALVGDARIVLIGSSH 277
Db 1375 LDGTVGAEVATRPAPAAATDRDPYIVGMACRYPGVASPEELWELLATGR----- 1425
Qy 278 GTHFYGARAMTQWLIKEKGFGVAARADWPDVAVRVVRYRGGE--TNADEALSGFE 335
Db 1426 -----DAVALDPD-----RGMDLDLVSADPDSSG-- 1451
Qy 336 RFPAMMNTVVRDFEVLRTNRORYESGALROAFYG-----LDYSLHRSIQEVL-S 388
Db 1452 -----TSVRSRGFFYDAGEF-DADFGISPRBALAMPQGRLLLEVAME 1495
Qy 389 YLDKVDPRAPARARVACPDHACADGQAYGFAPAAFGAGPSCREAVEQLV---DYQR 444
Db 1496 TVERAGVPAASIKSQTCGVFVGAAA--QGYG---TGAGQAAGSGEGYFLTGAAGSVVS 1548
Qy 445 NALAY-----ARQDGLADELFEYAQNNQTVRDAEVYRRAMPSSGRVTSWNLRDQMA 497
Db 1549 GRLSYTTGLEGPAVTVDTACSSSLVALHLLAQALRSGECS--SLALAGVT-----VM 1598
Qy 498 QTLGSLTLTHDRHLDAFPARIVVMAHNSHVDADATEVMADGQLTGQIVERYRGDESR- 556
Db 1599 ATPGIFVEFSRQGLAADGRCKAF-----DAADGTGMSG--VGMLLERLSIDARRN 1649
Qy 557 -----SIGRSTYTGTVTAASEWGIQKRAVRPAL-----HGSV 590
Db 1650 GHRVLAVRGSVANVDGASNGILTAPN--GPSQORVIRALANAGLAASDVADAVEAIGTG 1706
Qy 591 BELHQATDSFLVARSDAEAPL-----DYVRLGARIGVVLPA 631
Db 1707 TSL-GDPEAQALATYGQORERPELLGSIKSNIGHTOSAAAGVAVIMVLAHMGALPR 1765
Qy 632 TEROSHYLVAPADQDFAMIHID-QTRALEPLEVTSRKIAGENP 674
Db 1766 T-----LHV---DQ--PSTHVDWSAGAVELLTEPAEKPGISRP 1798

RESULT 2
US-10-724-598-49
Sequence 49, Application US/10724598
Publication No. US20060019352A1
GENERAL INFORMATION:
APPLICANT: BLANCHE, FRANCIS; CAMERON, BEATRICE; CROUZET,
JOEL; DEBUSCHE, LAURENT; LEVCY SCHILL, SOPHIE;
THIBAULT, DENIS
TITLE OF INVENTION: BIOSYNTHESIS OF COBALAMINS AND/OR COBAMIDES, DNA SEQUENCES
CODING FOR THESE POLYPEPTIDES, PREPARATION METHOD AND THEIR
US.
NUMBER OF SEQUENCES: 60
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN
STREET: 555 13TH STREET, N.W.
CITY: WASHINGTON
STATE: DISTRICT OF COLUMBIA
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10724,598
FILING DATE: 01-Dec-2003
CLASSIFICATION: 435
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/916,151
FILING DATE: 14-SEP-1992
APPLICATION NUMBER: PCT/FR91/00054
FILING DATE: 30-JAN-1991
ATTORNEY/AGENT INFORMATION:
NAME: F. P. CALVETTI
REGISTRATION NUMBER: 28,557
REFERENCE/DOCKET NUMBER: 1290-7213
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 857-7887
TELEFAX: (202) 857-7929
INFORMATION FOR SEQ ID NO: 49:
SEQUENCE CHARACTERISTICS:
LENGTH: 1275 amino acids
TYPE: Amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: Linear
MOLECULE TYPE: protein
HYPOTHETICAL: No
ORIGINAL SOURCE:
ORGANISM: Pseudomonas denitrificans
STRAIN: <Unknown>
INDIVIDUAL ISOLATE: <Unknown>
DEVELOPMENTAL STAGE: <Unknown>
HAPLOTYPE: <Unknown>
TISSUE TYPE: <Unknown>
CELL TYPE: <Unknown>
ORGANELLE: <Unknown>
FEATURE:
NAME/KEY: COBN
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION: Translation product of SEQ ID NO: 48
SEQUENCE DESCRIPTION: SEQ ID NO: 49:
US-10-724-598-49

Query Match 3.7%; Score 131; DB 6; Length 1275;
Best Local Similarity 20.0%; Pred. No. 0.11;
Matches 127; Conservative 87; Mismatches 235; Indels 186; Gaps 29;

Qy 65 LDAFVVRKLGAPGHDRAVALASGGRVYVNDVAVGRTTPQOLDIAERBEELLRE 124
Db 683 LDAYLCDLKEQMRIDGLHIFGVAPEGRLLT--DLTVALARVPRGLGEGGQSLQRAIAD 740
Qy 125 SAYRGSRPTDITGKTYIVVDGATGASWFAVQALRDAQP-----AQ 168
Db 741 AGLRGPAIFTSAGNRA-----RDAQFPDPLDQWSDTWGPKPS 780
Qy 169 IVIAVPAPESTGREFAGLVDDVCAITMPTFLAVGESFWDFQVTDSEVRLLATPTAG 228
Db 781 ILADLSDAFRTAGDIVERI-ELLANLVSGELACPDHMANTRAIVLGEIETRL-----K 833
Qy 229 PSLRPPASTAALAVLR---RVALDAPGVPTHEVLALVGDARIYVLIGSSHTGEFY- 283
Db 834 PSISNGAAMTGFLLTGLSGRFVAPGSGAPTR-----GRPDVLPPTGN-----FYS 880
Qy 284 -QARAANTQWLIKEKGFGVAARAE-----ADWPDAYRVRY---VRGLGEPDNNAD 328
Db 881 VDSRAVPTPAAYE---LGKKSALLRRYIQQDHGEFSSGGLTAKGTAMRTGGDDIAQA 937
Qy 329 EALSGERFPAMMNTVVRADFVEMWLRTRNORYE-----SGALROAGFYGLDL 376
Db 938 LALIGAK--PTW-----DMSRRVMGYEIVPLAVLGRPRVDVTLRISGFF-RDA 983
Qy 377 YS-----LHRSIQEVSYLDKVPRAAARARARVACPDHACADGQAYGFAP--AAGACP 429
Db 984 FPDQIALFDVAIRAVALBEDDADNMIAARRRARSRLAEVGVAARASRYRVFGAKP 1043
Qy 430 SCREAVEQLVD-----VORVALYARODGLADELFEYAQNNQTVR--- 472
Db 1044 GAYGAALQALIDKGGMETRADLAAYVTWCAIYAGGEBKARDLF--BERLRTIEAVV 1101

Qy 473 -----DAEYTYRAMESGRVTSMNLRDQHMAQTLSLTHLDRHLDAAPPARIYVW 521
Db 1102 QONDRREHDLSDSDYQ--FEQGMSA-----AAEQLGGRPAIYHNDHSRPEKPYIR 1152
Qy 522 AHNHVGED-ARATEV---MADGQLTLGQIVREYGDSEBSRIGFSTYGTVAASR-----W 573
Db 1153 SLEBEIGVVRARVNPRIIDGVM-----RHC-----YKCAFETIAITVDYMF 1194
Qy 574 GGIQRKAVRPALHSGVEELFHQTADSPFVSARLS 608
Db 1195 AFAATYTGAVR-----DHFEPAAYQAFIVDERVA 1222

RESULT 3
US-11-143-980-49
; Sequence 49, Application US/11143980
; Publication No. US20050272133A1
; GENERAL INFORMATION:
; APPLICANT: He, Min
; APPLICANT: Hucul, John
; APPLICANT: Haltli, Bradley A.
; APPLICANT: Wegenaar, Melissa M.
; APPLICANT: Graziani, Edmund
; APPLICANT: Summers, Mia
; APPLICANT: Kulowski, Kerry
; APPLICANT: Pong, Kevin
; TITLE OF INVENTION: Biosynthetic Gene Cluster for the Production of a Complex
; FILE REFERENCE: AM-101426US
; CURRENT APPLICATION NUMBER: US/11/143,980
; PRIOR FILING DATE: 2005-06-03
; PRIOR APPLICATION NUMBER: US 60/664,483
; PRIOR FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/576,895
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 49
; LENGTH: 7968
; TYPE: PR1
; ORGANISM: Streptomyces sp.
US-11-143-980-49

Query Match 3.6%; Score 126; DB 7; Length 7968;
Best Local Similarity 23.3%; Pred. No. 3.8;
Matches 186; Conservative 79; Mismatches 285; Indels 248; Gaps 45;

Qy 15 RRVFRDRRGRVLAELIAAYRDPDVIIVLGARGLFPVMEVA-AALHAPLDAFVVRKL 73
Db 7137 REVVRKADGARRHLHELTRDI-DLDAFVLFSSAGTVGNAGGSYAANAVALDGLAMRR- 7194
Qy 74 GAPG--HDEFVAGLASGRVNVNDVVRGIRITPQQLRDIAERREGLRLRESAYRGER 131
Db 7195 RABGLVATSVAMGMAESGMAA-----EMARSGQMD-PRSSALAAAGLV 7236
Qy 132 PPTDITGKTIVIVDDGLATGASMEFAVQ-----ALRDAQPAQ-IVIAVP 174
Db 7237 LAABET--TVNVADIDMATPFARFTASRPSLSELSDGSVSTEAADGEPADAFAETLE 7294
Qy 175 AAPSTCRFPAGLVDVVCATMPT-----PFLAVGESFMDPFQVTDDEVVR 220
Db 7295 AMAE---RERAAVTLDLVTRTHVAAVLGHATAEADIPARPFQBIG-----FDSLTAVELRN 7346
Qy 221 LLAPPTGAPSLRRPAA-----STADVLRVAIDA--PGGVPTHEVLAELVGDARIYVLIG 273
Db 7347 RL---TAATGVRFPAASYIVDYPTFPAALAEHVCREALGCGRTPAVVRPVDDEPIALIG 7403
Qy 274 ESSHGHTEFYOARAAMTQ--WLIBEKGFGAVA--AEADW-----PDA-YRVVRVYR- 319
Db 7404 MSCR-----PFGGVSSPPEDLWGLAEGRDVSDFPADGCMNLAEIYDDPDHPGSSYRA 7458
Qy 320 -GLGSDTNA-----DEALSGFER--FPAMMRNTVYVRDFV 351

Db 7459 GGFIDDAADAPDPCFPGTISPREBALAMPQQRLLLEVAMEAPERAMSPATL-KSRTGVFV 7517
Qy 352 EWLTRNRQRY---ESGALROA-GF-----YGLD-----LY 377
Db 7518 ---GTNGQDVAALASGAPRSGABGYLGTGSAASVASGRLATVFGLEGPAVTVDTACSSSLV 7574
Qy 378 SLRSTGIEVISYLDKVDPPRAAARARYACF-----DHACDDGQAVGFPA--FQAGP 429
Db 7575 ALHIAAQLRSGECSLALAGATVMAATPAAFLERSRRLAALAGRCFAAADGTGWC 7634
Qy 430 SCREAVTEQIYDVYORN---ALAYAR-----ODGLAEDELFYAQONQOTRDAEYVRAM 481
Db 7635 GVGMLIVERLSDERKNGRHYLVAVRGSAYVODG--ASNGLTAPNGSQ-----QRTVQAL 7688
Qy 482 FSGRVTSMNLRDQHMAQTLSLTHLDRHLDAAPPARIYVWAHNHSHVDARATEVWADGOL 541
Db 7689 ANARLSA-----TDID-----VNEHGT--GTSIGDPTEA----- 7716
Qy 542 TLGQIVREYGDSEBSRIGFSTYGTV-----TAASEMGIAQRKAVRPALHSGVEELF 594
Db 7717 ---QALATYVG-QGRSQNKPLWLSVKSNIQHTQAAAGVAGVI--KVMAMRHGVLPRTL 7770
Qy 595 HQTADS-----FLVSARLSRDAEAPLDVVRGRALIGVYVLPATERQSHYLVPR 643
Db 7771 HVDSPPHVDMAARVLELVAREWPRTGAP-----RRAGVSSFGVSGTNAHYI---- 7819
Qy 644 ADQFDAMIHIDQTRALRP 661
Db 7820 VEGQPVVARDPDRSAREP 7837

RESULT 4
US-11-143-980-48
; Sequence 48, Application US/11143980
; Publication No. US20050272133A1
; GENERAL INFORMATION:
; APPLICANT: He, Min
; APPLICANT: Hucul, John
; APPLICANT: Haltli, Bradley A.
; APPLICANT: Wegenaar, Melissa M.
; APPLICANT: Graziani, Edmund
; APPLICANT: Summers, Mia
; APPLICANT: Kulowski, Kerry
; APPLICANT: Pong, Kevin
; TITLE OF INVENTION: Biosynthetic Gene Cluster for the Production of a Complex
; FILE REFERENCE: AM-101426US
; CURRENT APPLICATION NUMBER: US/11/143,980
; PRIOR FILING DATE: 2005-06-03
; PRIOR APPLICATION NUMBER: US 60/664,483
; PRIOR FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/576,895
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 48
; LENGTH: 7102
; TYPE: PR1
; ORGANISM: Streptomyces sp.
US-11-143-980-48

Query Match 3.5%; Score 122; DB 7; Length 7102;
Best Local Similarity 22.5%; Pred. No. 6.5;
Matches 135; Conservative 84; Mismatches 214; Indels 166; Gaps 34;

Qy 82 AVGALASGRVNVNDVVRGRLRITPQQLRDIAERBEGR--ELLRBSAYRGERPPTDITGK 139
Db 1210 ASGTVAAGGVAPRTEVVRVVSQPDV---VGEAYGHVLEVLRSVQAMVAD--DLAGE 1263
Qy 140 TVIVV-----DDGLA--TGASMEFAVQALRDQAQPIQIYVAPAPAEPCREPAGLVD- 189
Db 1264 RLVVVTRGAVDTGADVADVAGAAVGLVRSQSENPRLV-----LVDT 1307

QY 190 ---DVVCATMPPTPLAVGESFMDRQVTDVEVRLLATPTAGPSLR-PAASTAADVLRR 245
DB 1308 DDLDGVSLP-GMLAT-----DEEQIVRSQAVVPRLARVPAGVSGRGS 1355
QY 246 VAIDAPGVPTHEVLAELV-----GDARIVLI-GESSHGTHEFYQARAAMTQWLI 294
DB 1356 GAVLVLTGCT---GVUGGLVSHNLVARRGVSRVLVLSRRGAAEGAAELRES-----L 1404
QY 295 EEKFGAANAADWPDAYRVNRVYRGLEDNADDEALSGERFPAMWMTVVDPEWML 354
DB 1405 EAGAENVIAACDAADREALAGVLSGL---SADFALSQ-----VV----- 1441
QY 355 RTRNQRRESGLRQAGFYGLDLYSLHRSIQEVISTLDKVDPRRAARARACDHAACAD 414
DB 1442 ---HAAGVLDLDD---GL-LTSLTR-----EEVEPVLRKVDAAANLHELTTGM 1481
QY 415 DGOAYG-FAAFAFG---AGPSCEBAVEQLVDVQNALAVARODGLAEDELFYAQQNAQ 469
DB 1482 DLSAFVLPSSAAGILGNAGGQ-----SYAANGFL---DALAAHRRAR 1521
QY 470 TVRDAEYVY---RAMEFSGRVTSMNLRDOH-MAQTIGSLTTHLDRHLDAEPARIYVMAH 523
DB 1522 GLPAVSIAMGFWEARSELTOHLSADDLARAHAVPMPTSQALDLPDATLADEPVLAAPL 1581
QY 524 NSHVGDARATEVMD-GQL--TLGQIVREKYGDESRKIG---FSTYTGTVTA--SEWG 574
DB 1582 N-----PQAWSDAGHLPVLRDLVRPRIRRAAETTGAPESASALGHRLLAAVDRBMD 1633
QY 575 GIAOKKVRPALHGSVEELFHQTADSFLVSARLSRDAEPADV-VRLGRAIGVYVLPAT 632
DB 1634 QVV-RELVNRHIAAVLRHAGSESVDTSRTFQETGFDLSITAVELNRRISAATG-VRLPAT 1690

RESULT 5

US-10-858-730-89
; Sequence 89, Application US/10858730
; Publication No. US200505556681
; GENERAL INFORMATION:
; APPLICANT: Bailey, Richard B.
; APPLICANT: Blomquist, Paul
; APPLICANT: Doten, Reed
; APPLICANT: Driggers, Edward M.
; APPLICANT: Madden, Kevin T.
; APPLICANT: O'Leary, Jessica
; APPLICANT: O'Toole, George
; APPLICANT: Trueheart, Joshua
; APPLICANT: Walbridge, Michael J.
; APPLICANT: Yorgey, Peter S.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR AMINO ACID
; FILE REFERENCE: 14184-030001
; CURRENT APPLICATION NUMBER: US/10/858, 730
; PRIOR FILING DATE: 2004-06-01
; PRIOR APPLICATION NUMBER: US 60/475, 000
; PRIOR FILING DATE: 2003-05-30
; PRIOR APPLICATION NUMBER: US 60/551, 860
; NUMBER OF SEQ ID NOS: 364
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 89
; LENGTH: 528
; TYPE: PRT
; ORGANISM: Mycobacterium leprae
US-10-858-730-89

Query Match 3.5%; Score 121.5; DB 6; Length 528;
Best Local Similarity 22.6%; Pred. No. 0.17; Indels 183; Gaps 27;
Matches 118; Conservative 48; Mismatches 174;

QY 30 ELAAAYDQPDVIVLGLARGGLPVAMEVAAALHAFLPAFVYRKUGAGDHFAVGALASG 89
DB 36 KLAAAV---PEADAL-LVRSATTVDAEVLAA--APKLIYAR--AGVGLDNVDVDAATAR 87

QY 90 GRVYN-----DQVVRGLRITPQQLRDIAREG--RELRRESAYRGRPEPTDITGKT 140
DB 88 GVLVNAPTSINISAAEHALALLAASRQLEADASLRHAHMKRSSFSG---TILFGKT 143
QY 141 VIVVDG-----LATGSMFAAVQALRDAQPAQ-----IVIAV 173
DB 144 VGVVGLGRICQLVAARIIAFGAHVIAVDPVVAPARAQIGIELMSFDLLARADFISVHL 203
QY 174 PAPESTCEBFAGLVD-DVVCATMPPTPL-----AVGESFW 208
DB 204 PKTPET-----AGLIDKEALAKTKPGVITVNAARGLVDEVALADAVRSGHVAALADLVF 258
QY 209 DFGQVTD---EVRRLATPTAGPSLRRAASTAADVLRR-----AIDAPGV 254
DB 259 ATPCCTDSPLELSQVVVTPHLGASTAEADDRAGTVASVRLALAGEFVPDVAVNDGV 318
QY 255 PTHEVLAELVDGARIVLIGSSHGTHEFYQARAAMTQW-LIEKFGAANAADWPDAY 312
DB 319 VNEEV-----APWLDVYKGLVVAALSDDELPAUL 348
QY 313 RVNRVYRG--LGEDTN-ADDEALSGERFPAMWMTVVDPEWMLRTNORYESGALRQ 368
DB 349 SV--HYRGELASBDEILRLSLRGL-----FSTVIEDAVTFVNAPALAAERGSAB 398
QY 369 --AGFYGLDLYSLHRSIQEVISTLDKVDPRRAARARACFDHACADGQAYGFAAFG 426
DB 399 ITTGSSEPN---HRSVVDV-----RAVASDGSVVNIAGTV- 430
QY 427 AGPSCEBAVEQLVDVQ-NV-----ALAYARQDGLAE 458
DB 431 SGF-----QLVQKIVQVNGNRPDLRAQGMNLVIRYVQDPAALGK 469

RESULT 6

US-11-205-109-15
; Sequence 15, Application US/11205109
; Publication No. US20050287641A1
; GENERAL INFORMATION:
; APPLICANT: Farnet, Chris
; APPLICANT: Zazopoulos, Emmanuel
; APPLICANT: Staffa, Alfredo
; TITLE OF INVENTION: GENE CLUSTER FOR RAMOPLANIN BIOSYNTHESIS
; FILE REFERENCE: 3002-205
; CURRENT APPLICATION NUMBER: US/11/205, 109
; PRIOR FILING DATE: 2005-08-17
; PRIOR APPLICATION NUMBER: US/09/976, 059
; PRIOR FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: US 60/239, 924
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 15
; LENGTH: 8695
; TYPE: PRT
; ORGANISM: Actinoplanes sp.
US-11-205-109-15

Query Match 3.3%; Score 115.5; DB 7; Length 8695;
Best Local Similarity 21.7%; Pred. No. 27;
Matches 193; Conservative 65; Mismatches 294; Indels 339; Gaps 46;

QY 31 LLAAYDQPDVIVLGLARGGLPVAMEVAAALHAFLPAFVYRKUGA----- 75
DB 7660 LLDVHRDG-----VGLVROERPLDGEVVARLRA-----VARRIGVSPATVMVAMARVLG 7709
QY 76 --PGHDEFVAGLASGRVNVDDVVRG--LKITPQQLR-----DIAE--REGRELL 121
DB 7710 VISGRDVFVFTLL-GRFSTGADRVGPPIITLPPRALIGTGGAATAVAEMRRLAELL 7768
QY 122 RRE-----SAYRGERP-----PTDITGKT----- 140
DB 7769 EHEHAPLTTAQAASGLSGULPLFTALFVNRHNTSPGADSPPAAGPTEGIRPVSMERTYV 7828


```

Qy 141 ---VTVDDGATGASMPAAVQALRDAQPAQIVAVPAAPESTCREPAGLVD----- 189
Db 7829 PISVAVDDGEGGIG-----VAVNAIPVYFPAVCELVATATSTLSALELFDGSGPDPAVG 7884
Qy 190 ---DVV-----CATMTPTPLAVGESE--WDFQV-- 213
Db 7885 ELVDLPGEGRSRLLEWMDTKARPVSESVPLAFERVAADAAVAVGEGVSWISREIDR 7944
Qy 214 -TDEEVRLATPTAGPSLRPP--AASTADVLR-VAIDAPGV-----PTHEVL 260
Db 7945 RSDVLASIVA---AGVGLSEPPVVALERSADVLAFLAVAKAGSVFVPVDSIPQIRID 8001
Qy 261 AELVGDAIVLIGSSSGHTHEFYQARAAMTOMLIEKGFCAVAAADMP-----DAYRV 314
Db 8002 A-VIADSRPVLVDS-----VDLPAAEADLPVPAAGAVYRM 8037
Qy 315 NRYVRG-----IGEDT-----MDEAL--SGFERF----- 337
Db 8038 --YTSGSTGRPKGVVTHQNLVDLATPTCWGSTPRVLPHAPHADASSYEIMVPLNGT 8095
Qy 338 ----PAMMMRTVVRDFVEMILRTNORYESGALRQAGFYGLDLSLRSIQEIVSYLDKV 393
Db 8096 VVVAPRRSIDATVARDLVGRHELTHVHTAGLAV-----LD-PSCFAGLTVETLGGDAV 8149
Qy 394 DPA--AARAPARYA-----CFDHACADGQAYGFAAFGAGPSCREAVEQ 438
Db 8150 SAEAVRBYKEANPGLRVQLYGPTETVLCATOHLLDQ-----VPIGRPLDTRVYV 8201
Qy 439 LVVDQR-----ALAYARODGLAE-----DELFAQOMACTVR 472
Db 8202 LDDLRFPVPGVVELVYAGSGLAGYAGMPGLTAERFVADPFSVGSLYXTGDLVMTD 8261
Qy 473 DAAYYRAMPFSGRTVSNLDOHMAQTGLSLTLDRHLDAAPARIYVMANSHVAGARA 532
Db 8262 DGVH----PAGRAD--QVKRGYRVERGEVATLAQHPVDSQVAVV--REDAPGDKRL 8314
Qy 533 TEVWADQQLTGQIVREHYGDESSIGFSTYGTV-----TA-----ASEWGG 575
Db 8315 VAVYVGDV-----EAYAQE--RLPGYVPSAFVHLALPLTANQKVDRAALPAPERET 8366
Qy 576 IAQRKAVRALHGSVEELFHOTADSFLVSARLSRDAEPD-----VRL 620
Db 8367 TTPGKAAPGPGNLIESMCA-----FAEVLGLDSVGPDDDFPAGGSHLALVALVQL 8421
Qy 621 ---GRAIGVYVLPATEROSHYL-----HVRPADQPMIHIQTRALEPL 662
Db 8422 KARQAVTVQDIMAAPVYSELMSLSMSIR--DSLGLTLPIRTGELPL 8470

RESULT 7
US-11-055-822-898
; Sequence 898, Application US/11055822
; Publication No. US20050260707A1
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Krieger, Burkhard
; APPLICANT: Schröder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauser, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING
; FILE REFERENCE: BGI-121ICPN
; CURRENT APPLICATION NUMBER: US/11/055,822
; PRIORITY FILING DATE: 2005-02-11
; PRIORITY APPLICATION NUMBER: 09/606,740
; PRIORITY FILING DATE: 2000-06-23
; PRIORITY APPLICATION NUMBER: 60/141,031
; PRIORITY FILING DATE: 1999-06-25
; PRIORITY APPLICATION NUMBER: 60/142,101
; PRIORITY FILING DATE: 1999-07-02
; PRIORITY APPLICATION NUMBER: 60/148,613
; PRIORITY FILING DATE: 1999-08-12
; PRIORITY APPLICATION NUMBER: 60/187,970

```

```

; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: DE 19930476.9
; PRIOR FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: DE 19931415.2
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931418.7
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931419.5
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931420.9
; PRIOR FILING DATE: 1999-07-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1158
; SEQ ID NO 898
; LENGTH: 449
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-11-055-822-898

Query Match      3.2%; Score 113; DB 7; Length 449;
Best Local Similarity 22.9%; Pred. No. 0.58;
Matches 63; Conservative 32; Mismatches 90; Indels 90; Gaps 11;

Qy 15 RRVPRRRRGRVLAELIAYRDQPDVIVGLARGGLPVAWVAAALHAFLDAFVVRKLG 74
Db 231 RVNVEARLEIGRKIAEAPAVGD---LVIPRESGTPAVVGFPAQASGIFPGGMVKN-- 284
Qy 75 APGHDEPAVGALASGGRVVN--DDVVR--GLRITPOOLRDIARREGRELLRBSAYRGER 131
Db 285 -----HYVGTFIQPSDTLRQULRIKLANPRLV----- 313
Qy 132 PPTDTGKTVIVDDGATGASMPAAVQALRDAQPAQIVAVPAAPESTCREPAGLVD 190
Db 314 ---IAGKRLVVDDSDIVRNTQRAVIRMLREAGAAEVHVRASPVKWC--FYGI--- 364
Qy 191 VVCATMPTPLAVGESWMDROYTDEVRRLATPTTAGPSLRPPASTADVLRRAIDA 250
Db 365 -----DPAITGELIAN-AVTSNDAAEVAEVRSA-----IGADTLGYVSDS 405
Qy 251 -----PGVPTHEVIAELV 264
Db 406 MVAATEQPANELCTIACFDGKTPMGLPQGNMADLV 440

RESULT 8
US-11-075-185-7
; Sequence 7, Application US/11075185
; Publication No. US20050266434A1
; GENERAL INFORMATION:
; APPLICANT: REEVES, CHRISTOPHER D
; APPLICANT: JULIEN, BRYAN
; APPLICANT: REID, RALPH
; TITLE OF INVENTION: BIOSYNTHETIC GENE CLUSTER FOR AMBUTICINS
; FILE REFERENCE: 010099.03
; CURRENT APPLICATION NUMBER: US/11/075,185
; PRIORITY FILING DATE: 2005-03-07
; PRIORITY APPLICATION NUMBER: US 60/551,103
; PRIORITY FILING DATE: 2004-03-08
; PRIORITY APPLICATION NUMBER: US 60/568,290
; PRIORITY FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: Patentin version 3.3
; SEQ ID NO 7
; LENGTH: 3507
; TYPE: PRT
; ORGANISM: Sorangium cellulosum
US-11-075-185-7

Query Match      3.2%; Score 111; DB 7; Length 3507;
Best Local Similarity 21.5%; Pred. No. 16;
Matches 177; Conservative 79; Mismatches 244; Indels 322; Gaps 48;
Qy 37 DQPDVIVLGLARGGLPVAWVAAALHAFLDAFVVRKLGAPGHDEPAVGALASGGRVVND 96

```

Db 626 DRVDVV-----QPALFMAVGLSAVMRAMGIEPSAVVSHSGEVAATAAVALTL-E 676
Qy 97 DVVNGELITPQQLDIA-----BREGRELLRESAYRGERPTDIT-GKTVI----- 142
Db 677 DAAAVVLRSQLVRRIRIAGGAMAVIERPVGEVEORLSRFGQLSVAAVNTGSTIVVSGDA 736
Qy 143 -VVDDGLA-----TGASMFVAVALRDAPQIVIAVPAPESTCEBAGL 187
Db 737 AAVDRLLAELEHEBEPARVVVVDYASHAHVDLILPELEACLAVERPA----- 785
Qy 188 VDDVYCA-----TTPTEPLA---VGESFW--DFRQVT--DEEVRLLA-----TTPAG 228
Db 786 -----CA1PLVSTYTGVELAGPELGAAYWCNRLREPRALDRLADLADGHVEVEVSAH 840
Qy 229 PSLRPAASTADVLRRVAIDAPGCVPTHEVLAEVLGDARIVLIGESSH----- 277
Db 841 PVLAIPLTAASAE-----RGGVVVGSIQDDGGLGRVLSALGALHVOGSHVEMARV 891
Qy 278 -----GTHEFYQARAAMTOWL-----IEBKFG-----AVAAEADWPD 311
Db 892 LAPYGNLVLDLPTAFOQR-----YMLEASRSRIDASDLGAATGRPLGAATRVAGTDS 947
Qy 312 YRVNRYRIGGEDTNADEA-LSG---FER--FPAAMMRNTVVDVPEWLTTRNQRYESGA 365
Db 948 YIL-----AGRLSTABHPWLSGVVPERTLFPA-----TGFLER-----A 982
Qy 366 LRQAGFYGL-----DLY-----SLHRSIOEVI SYLDKVPRA-AAARARAYACFD 409
Db 983 LEAADAMGAVAGVTELVVPAPILPARGAVH--VQVAVQGPDEAGRRFPFSYSRAETGGLD 1040
Qy 410 -----HACADGQAYGFPA-----FGAGPSCEREAVEQLVDVQRNALAYARODGL 455
Db 1041 AEWTLHATGILLGARASAAADTGLEAMPREGAAL-----VDVSD--AYAR----- 1083
Qy 456 LAEBELFYAQNACTVR---DAEYVYRAMPSGAVTS-----WMLRQDMA-QTLGS 502
Db 1084 LEDGAVYAPSLRALVBAWQAEERRIYARAVLPFGATQGHGLPMLWPAALHMLLVVLGQ 1143
Qy 503 LETHLDBHLDPARPIV--W-----AHNSHVGADARATEVMAD 538
Db 1144 -----DHAHGVLLPRMSDVTLLAAGATLRVRVELADDAEHV-SASLTMAAD 1193
Qy 539 GQ--LTLGQIVREB-----YGDSESRISGFS--TYTGTVTAASEWGIA 577
Db 1194 GQPAVATGSGVEVRATAAQAAMSTATQHLVGVEMKAVALLAEPRASGEQVLGPDELA 1253
Qy 578 QRAVPRALHSGVEELFHQTADS-----FLVSARLSRDAEAPLDVVRIGRAIGVY 628
Db 1254 TRUGARRA--GNIDELF---ADGEAARPARRLVVDARTTRDGDVP----- 1294
Qy 629 LPATERQSHLVHVRADQPDAMIHIDQTRALEPLEVTSRWIA 670
Db 1295 -----AAVH-QATR-QALELVQRMLA 1313

RESULT 9
US-11-075-185-4
; Sequence 4, Application US/11075185
; Publication No. US20050266434A1
; GENERAL INFORMATION:
; APPLICANT: REEVES, CHRISTOPHER D
; APPLICANT: JULIEN, BRYAN
; APPLICANT: REID, RALPH
; TITLE OF INVENTION: BIOSYNTHETIC GENE CLUSTER FOR AMBRUTICINS
; FILE REFERENCE: 010099.03
; CURRENT APPLICATION NUMBER: US/11/075,185
; PRIOR APPLICATION NUMBER: 2005-03-07
; PRIOR APPLICATION NUMBER: US 60/551,103
; PRIOR FILING DATE: 2004-03-08
; PRIOR APPLICATION NUMBER: US 60/566,290
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 61

; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 4
; LENGTH: 3689
; TYPE: PRT
; ORGANISM: Sorangium cellulosum
US-11-075-185-4

Query Match 3.2%; Score 111; DB 7; Length 3689;
Best Local Similarity 22.0%; Pred. No. 17;
Matches 177; Conservative 63; Mismatches 253; Indels 312; Gaps 41;

Qy 13 SPFRVRPD--RRAGVLAELLAAYRDPQVYVLGLARGGLPYAVMAAALHAPDAFV 70
Db 2774 APRRLVLDLTAADAAGAVVAHNAARD-----ALSLVQG-----WLAAEQLAATELVVVT 2823
Qy 71 RKUGARPHDE-----FAVGALASGGRVVVNDVVRGLRITPQO-----LRDIAEREG 117
Db 2824 RGAVAAPRDEGVAALGPAAVWGILRTRVHNADRTYRMIDLGSGADDMALLRRLTAABE 2883
Qy 118 RELLRRESAYRGERPPTDITGKT--VIYDDG-----LATGASMPA-----AVQALRD 163
Db 2884 PELALRAGGARA--PRUDAAGETDGLAPPDGARSRLSIRTGSEFDALHLADAPDALRP 2941
Qy 164 AQAQIVIAVPAPESTCEBAGLVDDVVCATMPTFPLAVGESFMDPROVTDSEVRLLA 223
Db 2942 LGEGQVRLAVRA-----TGL-----NFR-----DVANVLG 2966
Qy 224 T--PTAGPSLRRAASTAADVLRRVAIDAAG-----GVPTHEVLAEVLGDARIVLI 272
Db 2967 TYRGEAGP-LGLEGAGVLDVGGVATALRQGDVWGILHAGMATHAVV-----DARLL-- 3018
Qy 273 GESSHGTH-----EPYQAR--AAATQWLEBKFGAV----- 302
Db 3019 -----THIRGLSFVEATIPPAFLTALYGLDLTAKAGQRLVHAAAAGVGMADVQV 3072
Qy 303 -----AAADWPDAYRVNRYVRIGLGEDTNADEALSGEERPPAMMRNTVVRDPV 351
Db 3073 ARLMGAEVFATAGEKMP-----ALRKGMD-QAHIISSRTLHR-----KATL 3115
Qy 352 EMLRTNORYESGALRQAG-FYGLDLSYLSHRSIOEV-ISYLDKVPDRAAARAR--ARYAC 407
Db 3116 D--ATRGQGVDDVLDLAGEFVNDASLDLPRGRFVEMKSDVRDEBRAKHPGRYTA 3173
Qy 408 FDHACADGQAYGFPAAFGAGPSCEREAVEQLVDVQRNALAYARQGLLAEBELFYAQN 467
Db 3174 FD-----LIDAGPDHIOAMLRRLVPL-----FEBGVLA--PLPFA-- 3206
Qy 468 AQTVRDAEVYVYRAMPSGRVTSWMLRDQMAQTL-----GSLTHLD 508
Db 3207 VHDLRBAHPAHRMANAR-----HVQKVLVPPALDDPGTALITGGTGLGRQIA 3257
Qy 509 RHL-----DAPPARIVWA----- 522
Db 3258 RHLVAHGVRLVLTTSRRGNDAPDAALVGSLLAAGATVEVAACVTRBDALAAVQAI 3317
Qy 523 -----HNSHVD-----ARATEVMADGQTLTGQIVREB-----YGD 553
Db 3318 PAARPLTAIVHTAAVDDGIVAGLSAEQLARVLPRKVDGAWRLYEATRTDAPLAFMLFSS 3377
Qy 554 ESRSIGFSTYTGVTVAASEWGIAQRKAVR--PAL-----HGSVEELFHQTADSFLV 603
Db 3378 VACTLGSSGQANVAANAFLDGLAELRTRGVPAWSLANGFWEQGGIGMTHIGAADL-- 3435
Qy 604 SARLSRDAEAPLDVVR-----LGRAI 624
Db 3436 -ARLKGQGIAPMTVAAGRLDLRAL 3459

RESULT 10
US-11-143-980-47
; Sequence 47, Application US/11143980
; Publication No. US20050272133A1
; GENERAL INFORMATION:

APPLICANT: He, Min
APPLICANT: Hucul, John
APPLICANT: Haultli, Bradley A.
APPLICANT: Wagenaar, Melissa M.
APPLICANT: Graziani, Edmund
APPLICANT: Summers, Mia
APPLICANT: Kuliowski, Kerry
APPLICANT: Pong, Kevin
TITLE OF INVENTION: Biosynthetic Gene Cluster for the Production of a Complex
FILE REFERENCE: AM-101426US
CURRENT FILING DATE: 2005-06-03
PRIORITY FILING DATE: 2005-06-03
PRIORITY FILING DATE: 2005-03-23
PRIORITY FILING DATE: 2004-06-03
NUMBER OF SEQ ID NOS: 72
SOFTWARE: PatentIn version 3.3
SEQ ID NO 47
LENGTH: 5712
TYPE: PR1
ORGANISM: Streptomyces sp.
US-11-143-980-47

Query Match 3.1%; Score 109.5; DB 7; Length 5712;
Best Local Similarity 20.1%; Pred. No. 42;
Matches 157; Conservative 86; Mismatches 265; Indels 275; Gaps 42;

19 RDREAGVLAELAAAYDQPVYVGLAR---GGLPYAME--VAALALPLDAFVVRK- 72
864 RDCUTGPGTLVGLTRDRPEFOALVTALAEVSGVEVAMSPVSGGRRIPLPTAFQRO 923
73 ---LGAPGHDFPAVGALSGRRVVDVVRGLRTTPQQLDIAREGELLRESAARG 129
924 RYMFSAPEPE---SG-----TTPGHGVT-----SGRE--KDTGTGSG 955
130 ERPEPTDITGKTVIVDDGATGSMFAAVQALRDAQPAQIVYAPAA---PESTCREFA 185
956 DEAF-----DTG-PSGGETLGMVRA-----HAAVVLGYASATAIGAEHTFQKG 998
186 GLVDVVCATMPPTPLAVGESFMDPROVTDSEVRRLATPTAGP-----SLRPPAAT 238
999 -----FDSITAVELCERLCAATLPLPGLTLPDYTPPA-L 1033
239 AADVLRV---AIDAGCVTHEVLAELVGDARIVLIGSSHCTHEFYQARAAMTQWL 293
1034 AEHLHRRLLHGRTRDQAAAPATVPTPD-----GGDPVIVVGSCR-----FGRASPE-- 1080
294 IEKGFCAVAABADWPDAYRVNRVYRGLGEDTNADEALSGFERPPAW---MMRTVVRD 349
1081 -----DLWRI-----VADGEDALSGFSPDRGMWLAGLYHPDRPH 1115
350 FVEWLRITNQRYESGALRQAGFYGLDLYSLRHSIOEVI SYLDKVDPPRAABARARYACFD 409
1116 GTSYARDCGFLYDAEF--DAGFFGI-----SPREA---EAMPDQORLLLETSMWALE 1163
410 HACADDCGAYFAAAGFAGPSCERAEVQOLYDVQRMALAYARQGLAEDELFTAAQQAQ 469
1164 RA-----GIPAEHIGSSSTG-----VFIGASSGVYADAGEAEG---YOUTGTAA 1206
470 TVRDAEYVYRAMFSQ-----RVTSMNLRDQMA-QTIG---SILLTHDRHLDAFPARI 518
1207 SVASGRVSYTIGLEGPATVYDTACSSSLVALHIAVQSLRAGECSTALAGCTVWATPAMF 1266
519 VVAHNSHVQ--DAPATEV-----WADGQTLTGQIVREERYGDESR-----S 557
1267 VEFSGRGLAMDGRCFAAADAAGTGAEG---VGLVVERLSDERRGHVLAVERGSA 1323
558 IGFSTYTGTVTAASWGGIAQKAVRPL-----HGSVELEFHQ--TADSF 601
1324 VNODGASNGLTAPN---GPSQORVITRQALASGLVASDVDAVEAHGTGTGTDPIEAQAL 1380

602 LVASRLSRDAEAP-----DVARLGAIGGVVLPATEROSHLYHR 642
1381 LATYGGCRDADRPLMLGVSXNSIGHTQAAAGVAGVIVVMAVRGVLPRT-----LHVD 1434
643 PADQDAMIHD-----QTRALEPLVTSRMWINGEN-----PPT-Y 677
1435 -----EPSTHVDMSGRRVELLTGTPWPTTGGLRACVSSFGVGTAAHYLLEQVETAR 1489
678 PTG 680
1490 PTG 1492

RESULT 11
US-11-075-185-8
Sequence 8, Application US/11075185
Publication No. US20050266434A1
GENERAL INFORMATION:
APPLICANT: REEVES, CHRISTOPHER D
APPLICANT: JULIEN, BRYAN
TITLE OF INVENTION: BIOSYNTHETIC GENE CLUSTER FOR AMBRUTICINS
FILE REFERENCE: 010099.03
CURRENT FILING DATE: 2005-03-07
PRIORITY FILING DATE: 2005-03-07
PRIORITY FILING DATE: 2004-03-08
PRIORITY FILING DATE: 2004-03-08
PRIORITY FILING DATE: 2004-05-04
NUMBER OF SEQ ID NOS: 61
SOFTWARE: PatentIn version 3.3
SEQ ID NO 8
LENGTH: 2197
TYPE: PR1
ORGANISM: Sorangium cellulosum
US-11-075-185-8

Query Match 3.1%; Score 109; DB 7; Length 2197;
Best Local Similarity 21.9%; Pred. No. 12;
Matches 95; Conservative 47; Mismatches 159; Indels 132; Gaps 21;

2 LMTAADVTRSPRRVRD--RREAGVLAELAAAYDQPVYVGLARGLPYAMEVAAA 60
1630 LETVAACARCAIGVSVGLHLDGTGLVS---ARLRKSPVVYLD-----EAIAG 1676
61 LHAFLDAFVVRKLGAPG---HDEFAV---GALASGRVVDVVRGLRI-----TPQ 107
1677 FELFPGVCAIRRFHESGVMTTRQPEEFARVVGSMAGPALTVBENILRRFRAVVTNDTTA 1736
108 QLRDIA--EREGBELR-----RESAYRGEPPTDITGKTVIVVD 145
1737 NLRALAVDQRTKRAHNSYVNVLESLEDAFGLAGRQHRHADRGVEIERD--DSSARVIN 1795
146 DGLATGSMFAAVQALRDAQPAQIVYAPAAPESTCEFPAGLVDDVVCATMPPTPLAVGB 205
1796 LYVITSAS-----FRGHTGSEIAGSVLGTHTDTRDYMDLERRIPRET----- 1838
206 SFMDFRQVTDSEVRRLATPTAGSRLRPAASTADVLRVAIDAPGCVTHEVLAELVQ 265
1839 ---DFGRV-----FPAGP-----ATAVEIAVNLGL-----LAARKG 1867
266 DARIVLIGSSHCTHEFYQARAAMTQWLEEKGFCAVAABADWPDAYRVNRVYRGLGEDT 325
1868 -SALLVLGSG-----PFTGLGLVSHAEBSPLDALVESCPMSKVIAVDPFGGAAAE 1921
326 NA-----DEALSGFERPPAMMRNTVVDVFEWLRITNQRYESGALRQAGFYGLDLYSLR 381
1922 EAKUTSDV--GF-----VMLETLQSD--W-----GGLRSVPDAVLEVIDRHR 1960
382 SIQEVISYLDKVD 394
1961 ---ERSGYLVGD 1970

RESULT 12
US-11-096-568A-20149
; Sequence 20149, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 20149
; LENGTH: 829
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(829)
; OTHER INFORMATION: Ceres Seq. ID no. 12377953
US-11-096-568A-20149

Query Match 3.1%; Score 107.5; DB 7; Length 829;
Best Local Similarity 21.2%; Pred. No. 3.7;
Matches 125; Conservative 70; Mismatches 207; Indels 187; Gaps 31;

QY 142 IVVDDGLATGASMPFAVQALR-----DAQPAQIVIAVPAPESTC-----REFAGLVDDV 191
DB 1 MWLSDGHTMOSMLATBNARVRDGSIOKSIYHL-----ECTCSITQSRIITIVIKDV 56
QY 192 V---CATM--PTPEFLAVGESFWDPROVTDSEVRLLA-----TPTAGPS-LRRPAS 237
DB 57 LQSECALIGNPRPY-----EMRNQNECGTNI PANAAQANTGYSSGGLGSPAS 108
QY 238 TAADYLRRVAIDAFCG-----VPTHEVLAEVGD-ARIVLIG-ESSHGTHEFYQARAAM 289
DB 109 RPAQVANNVPGSSGYGHTGVPPPIAPAVESVPNVASGVYGTTSAHNT-----MNAGM 163
QY 290 TQWLIIEKGFCAVAEADMPDAYRVNRYVRLGE--DTNADEALSGFERFPAMMRTTVV 347
DB 164 TQSNLQGRSLNSHPNQ-----RFAVPSMAGSGAPGNTYQGPASFTQGPFGHMRTPV 218
QY 348 RDEFWMLRTNRQYESGALRQAGFY-----GLDLYSLHRSIOEVI SYLDKV 393
DB 219 -----SKNDAYRLVPLAQINPYLDKWTIKVVTAKTDLRFYNNARGAKVFSF-DLL 269
QY 394 DPEPAAARARVACFD-----SKNDAYRLVPLAQINPYLDKWTIKVVTAKTDLRFYNNARGAKVFSF-DLL 409
DB 270 DEQ---RGEIRATCFNTQADQFNLIEVDKYYLISKSLKPAOKKFNLSLHVEISLDSR 326
QY 410 ---HACADG-----QAYGFAAFAFGAPSCERAVEQLVDVQRN---ALAYARQDGLAE 458
DB 327 TSIEVCADDDSNIPROQYNF-RKISEIENIEKAIIDLIGIVTSVGSVTFIRKQGV--- 382
QY 459 DELFYAQNAQTVDAEYVYRAMFSGR--VTSWNLRDQMAQTLGSL-THLDRLHLDAP 514
DB 383 -----ETQRLTLEKD-----MSGRSVQLTLMGKLCVAEGNQLOSLCDSGLNPVIALK 430
QY 515 PARIVVAHNSHVGDARATEVMADGQLTGQIVREYRGDSRSIGFSYTGTVTAAS--- 571
DB 431 GARVTVDSGRS-VSSAGSTQIKIDPEIPEASLRMYA-----TGKTPAACVSL 478
QY 572 ---EMGIAQRKA---VRPALHGSVEE-----LFHQTDASFLVSA 605
DB 479 SVVSMGTCVRKSIQIKDENLQLEKPDFTYKAAISHLIADNFCPPA 527

RESULT 13
US-11-096-568A-20148
; Sequence 20148, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:

; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 20148
; LENGTH: 843
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(843)
; OTHER INFORMATION: Ceres Seq. ID no. 12377952
US-11-096-568A-20148

Query Match 3.1%; Score 107.5; DB 7; Length 843;
Best Local Similarity 21.2%; Pred. No. 3.8;
Matches 125; Conservative 70; Mismatches 207; Indels 187; Gaps 31;

QY 142 IVVDDGLATGASMPFAVQALR-----DAQPAQIVIAVPAPESTC-----REFAGLVDDV 191
DB 15 MWLSDGHTMOSMLATBNARVRDGSIOKSIYHL-----ECTCSITQSRIITIVIKDV 70
QY 192 V---CATM--PTPEFLAVGESFWDPROVTDSEVRLLA-----TPTAGPS-LRRPAS 237
DB 71 LQSECALIGNPRPY-----EMRNQNECGTNI PANAAQANTGYSSGGLGSPAS 122
QY 238 TAADYLRRVAIDAFCG-----VPTHEVLAEVGD-ARIVLIG-ESSHGTHEFYQARAAM 289
DB 123 RPAQVANNVPGSSGYGHTGVPPPIAPAVESVPNVASGVYGTTSAHNT-----MNAGM 177
QY 290 TQWLIIEKGFCAVAEADMPDAYRVNRYVRLGE--DTNADEALSGFERFPAMMRTTVV 347
DB 178 TQSNLQGRSLNSHPNQ-----RFAVPSMAGSGAPGNTYQGPASFTQGPFGHMRTPV 232
QY 348 RDEFWMLRTNRQYESGALRQAGFY-----GLDLYSLHRSIOEVI SYLDKV 393
DB 233 -----SKNDAYRLVPLAQINPYLDKWTIKVVTAKTDLRFYNNARGAKVFSF-DLL 283
QY 394 DPEPAAARARVACFD-----SKNDAYRLVPLAQINPYLDKWTIKVVTAKTDLRFYNNARGAKVFSF-DLL 409
DB 284 DEQ---RGEIRATCFNTQADQFNLIEVDKYYLISKSLKPAOKKFNLSLHVEISLDSR 340
QY 410 ---HACADG-----QAYGFAAFAFGAPSCERAVEQLVDVQRN---ALAYARQDGLAE 458
DB 341 TSIEVCADDDSNIPROQYNF-RKISEIENIEKAIIDLIGIVTSVGSVTFIRKQGV--- 396
QY 459 DELFYAQNAQTVDAEYVYRAMFSGR--VTSWNLRDQMAQTLGSL-THLDRLHLDAP 514
DB 397 -----ETQRLTLEKD-----MSGRSVQLTLMGKLCVAEGNQLOSLCDSGLNPVIALK 444
QY 515 PARIVVAHNSHVGDARATEVMADGQLTGQIVREYRGDSRSIGFSYTGTVTAAS--- 571
DB 445 GARVTVDSGRS-VSSAGSTQIKIDPEIPEASLRMYA-----TGKTPAACVSL 492
QY 572 ---EMGIAQRKA---VRPALHGSVEE-----LFHQTDASFLVSA 605
DB 493 SVVSMGTCVRKSIQIKDENLQLEKPDFTYKAAISHLIADNFCPPA 541

RESULT 14
US-11-096-568A-20147
; Sequence 20147, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01

NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 20147
; LENGTH: 876
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(876)
; OTHER INFORMATION: Ceres Seq. ID no. 12377951
US-11-096-568A-20147

Query Match 3.1%; Score 107.5; DB 7; Length 876;
Best Local Similarity 21.2%; Pred. No. 4;
Matches 125; Conservative 70; Mismatches 207; Indels 187; Gaps 31;

QY 142 IVVDDGLATGASMFPAVOALR---DAOPAOIVLAVPAPESTC-----PEPAGLVDDV 191
DB 48 MVLSDGERTMOSMLATENAVRDGSIKGSIVHL---ECTGCTISRKIIIVIKLDV 103
QY 192 V---CATW--PTPLAVGESFWDPROVTEDEVRLLA-----TPTAGPS-LRRPAS 237
DB 104 LQSCAIGNRPY-----EMRNQNEGCTNIPANAQAQNGTSSGGLGSPAS 155
QY 238 TAAVLRVALDARCG-----VPTHEVLAVGD-AAIVLIG-ESSHGTIEFYQARPM 289
DB 156 RPAOVANNVPYSGSYGHTGVPEIAPAVESVPNVASGVYGTTSANHT-----MNAGM 210
QY 290 TQMLIEKGFQAVAAEADWPDAVRNRYVRGLGE--DTNADALSGEPFRPAMMMRNTVV 347
DB 211 TQSNLQQRSLNSHPNQ-----RFAVPSMAGSGAGANTYGGPAGSFYQPPGHMNTFV 265
QY 348 RDPFWEMLTRNQRYESGALROAGFY-----GLDLYSLHRSIOEIVISYLDKV 393
DB 266 -----SKNDARLVPLAQLNPYLDKWTIKVRYTAKTDLRFYNNAGAKKVSF-DLL 316
QY 394 DPRAPARARAYACD-----
DB 317 DEQ---RREIRATCFNTQADQFNLIEVDKYLLSKGSLKPAOKKENSINHEYEISLDSR 373
QY 410 ---HACADG-----QAYGFAAFGAGPSCEBAVEQVLDVQRN---ALAVARODGLLAE 458
DB 374 TSIEVCADDDSNIPROQYNF-RKISEINIEKDLIDLIGVTSVGPSTVTRKDGV--- 429
QY 459 DELFYAQOACTVDAEYVYRAMFSGR---VTSWNLARDQHAQTLGSL-LTHLDRLDAP 514
DB 430 ---ETQRTLELD-----MSGRSVQLTWKGLCVARNGNQSLCDSGLNPVLA-LK 477
QY 515 PARIVWAHNSHVGDARATEVWADGQLTLGQIVRERYGDESRISIGFTYGTVTAA-- 571
DB 478 GARVTDYSGRS-VESAGSTOLKIDPEIPEASLRWYA-----TGCKTAAACVSL 525
QY 572 ---EMGGAORKA---VRPALHGSVE-----LPHQTADSEFLVSA 605
DB 526 SVSMGTCVRKSIQIKDENLQLEKDPITVKAISHLADNFCYPA 574

RESULT 15

US-11-075-185-6
; Sequence 6, Application US/11075185
; Publication No. US20050266434A1
; GENERAL INFORMATION:
; APPLICANT: REEVES, CHRISTOPHER D
; APPLICANT: JULIEN, BRYAN
; APPLICANT: REID, RALPH
; TITLE OF INVENTION: BIOSYNTHETIC GENE CLUSTER FOR AMBRUTICINS
; FILE REFERENCE: 010099.03
; CURRENT APPLICATION NUMBER: US/11/075,185
; PRIOR FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: US 60/551,103
; PRIOR FILING DATE: 2004-03-08
; PRIOR APPLICATION NUMBER: US 60/568,290
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 61

SOFTWARE: PatentIn version 3.3
; SEQ ID NO 6
; LENGTH: 1892
; TYPE: PRT
; ORGANISM: Sorangium cellulosum
US-11-075-185-6

Query Match 3.1%; Score 107.5; DB 7; Length 1892;
Best Local Similarity 24.0%; Pred. No. 12;
Matches 131; Conservative 42; Mismatches 187; Indels 185; Gaps 27;

QY 2 LMTAAAVTRSPRRVRD--RREAGRVLAELAAVYDQDVIVLGLARGLPFAWEVA 59
DB 1286 LLEALAD-GAAPPRLVVDLTAAGAGAVVAVHAAVG---ALALVG---WLAAP 1334
QY 60 ALHAPLDAFVVRKUGARSHS-----FAVGALASGRVVVNDVVRGRLITPOQ--- 109
DB 1335 QLAATELVVTRCAVAVGPBGVDALGPAAVWGLLRTRTAEPRAVAVLDDVGRPLDGA 1394
QY 110 -----RDIAREGRELRRSAYRGERPPTD-----ITGKT----- 140
DB 1395 LRRALAAGTPELVSRSGEARAPRLREVRSSEPAAPATRLDPDGTALITGTGELGRH 1454
QY 141 -----
DB 1455 VAKHLVTAHGVRLVLTSSRGMDAPDAALVDELRAGAATVDVAACDADAAALAAVE 1514
QY 172 AVPA--PESTCRBPAGLVDDVCAITMPTPLAVGESFWDPROVTEDEVRLLATPTAGPS 230
DB 1515 AIFPARPLTAVHTAGVLDSDSVT-----KLSAEQLAKVL----- 1549
QY 231 LRRPAASTADVLRRAVIDAPGVPTHEVLAEVGDARIVLIGSSHGTIEFYQARAMT 290
DB 1550 -----RPKVD--GAFTLHE-LTKAPPLAFLF--SSAAGT----- 1580
QY 291 QMLIEKGFQAVAAEADWPDAVRNRYVRGLGSDTNADEALSGEPFRPAMMMRNTVV-- 347
DB 1581 ---LGSFGQAVYAAANTFLDALASHLRARGV-----PAMS-----LAWGFMAQTGLGM 1625
QY 348 ---RDPFWEMLR-----TRNQRYESGALROAGF-----YGLDLYSLHRSIOEIVISY 389
DB 1526 TAILGAADIDARKHGVSNPVAQGLRLIDRALAQATVLPALDSSIQRA-----GSN 1681
QY 390 LDKYDP--RAAARARARAYACFDHACDDGQAYGFAAFGAG---PSCBEAAVEQLVDVQ 443
DB 1682 AGVPPLRLRGLVRAPARRTAASAAGANGTG-AAALRLRLSPLPGAERQKV--LLDLV 1738
QY 444 RNALA 448
DB 1739 RTEIA 1743

Search completed: March 23, 2006, 06:29:34
Job time : 19.3423 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using SW model

Run on: March 23, 2006, 04:56:29 ; Search time 38.2848 Seconds
(without alignments)
1652.628 Million cell updates/sec

Title: US-10-617-038-24
Perfect score: 736
Sequence: 1 MATTLPVQRHPRSLPPEFSE.....SVAVSEKPTKEHIQIRSTN 144

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21:*

- 1: geneseqp1980s:*
- 2: geneseqp1990s:*
- 3: geneseqp2000s:*
- 4: geneseqp2001s:*
- 5: geneseqp2002s:*
- 6: geneseqp2003s:*
- 7: geneseqp2003bs:*
- 8: geneseqp2004s:*
- 9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	736	100.0	144	2	AAW94900 Alpha-cry
2	736	100.0	144	5	AAW50751 Mycobacte
3	736	100.0	144	8	ADI37303 M. tuberc
4	736	100.0	330	5	AAU74589 Antigenic
5	736	100.0	331	2	AAV32060 Mycobacte
6	736	100.0	331	7	ADF69755 Fusion pr
7	736	100.0	368	2	AAV32069 Mycobacte
8	736	100.0	368	5	AAU74598 Antigenic
9	736	100.0	368	7	ADF69775 Fusion pr
10	736	100.0	433	5	AAV32065 Mycobacte
11	736	100.0	433	5	AAU74594 Antigenic
12	736	100.0	433	7	ADF69765 Fusion pr
13	736	100.0	856	2	AAV32064 Mycobacte
14	736	100.0	856	7	ADF69763 Fusion pr
15	736	100.0	859	5	AAU74593 Antigenic
16	736	100.0	875	7	ADA26365 Mycobacte
17	718	97.6	143	9	AD247757 Mycobacte
18	356	48.4	131	2	AAK12489 Peptide d
19	245	33.3	48	2	AAK85672 12 KD M.t
20	245	33.3	48	2	AAK92889 Mycobacte
21	245	33.3	48	2	AAW18183 N-termina
22	245	33.3	48	2	AAW75570 M. tuberc
23	245	33.3	48	6	ABG74430 M. tuberc
24	245	33.3	48	6	ABU56334 M. tuberc

25	245	33.3	48	7	AAE39277 M. tuberc
26	245	33.3	48	7	ADF45119 M. tuberc
27	245	33.3	48	8	ADO36804 12KD majo
28	245	33.3	48	8	ADU64179 Mycobacte
29	164.5	22.4	159	6	ABP57444 Mycobacte
30	156	21.2	146	5	ABG77903 High leve
31	150.5	20.4	188	8	ADQ78294 S. aurant
32	148	20.1	164	9	AEK39411 L. pneumo
33	148	20.1	168	9	AEK35982 L. pneumo
34	146.5	19.9	151	9	ABM94473 M. xanthu
35	143	19.4	158	4	AAU63279 Propionib
36	143	19.4	158	6	ABM59798 Propionib
37	142	19.3	151	8	ADT60218 Plant pol
38	141.5	19.2	154	8	ADM48277 Polypept
39	134	18.2	170	8	ADX74462 Plant ful
40	133.5	18.1	142	7	ADC97715 Streptoco
41	133.5	18.1	154	7	ADC97714 Streptoco
42	133	18.1	151	7	ADP75373 Wheat SHS
43	133	18.1	151	7	ADL18503 Wheat pro
44	133	18.1	193	8	ADY10853 Plant ful
45	132	17.9	139	7	ADC94697 E. faeciu

ALIGNMENTS

RESULT 1
AAW94900
ID AAW94900 standard; protein: 144 AA.
XX
AC AAW94900;
XX
DT 11-MAY-1999 (first entry)
XX
DE Alpha-crystallin heat shock protein.
XX
KW Tuberculosis; Mycobacterium; alpha-crystallin heat shock protein;
KW acr gene; attenuated; vaccine; antimycobacterial; pathogenic.
OS Mycobacterium sp.
XX
FN W09902670-A1.
XX
PD 21-JAN-1999.
XX
PE 09-JUL-1998; 98MO-US014227.
XX
PR 10-JUL-1997; 97US-0052199P.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Barry CE, Yuan Y, Crane D;
XX
DR WPI; 1999-120871/10.
XX
N-PSDB; AAX17761.
XX
PT Novel attenuated strains of Mycobacterium tuberculosis and M. bovis -
PT useful in the production of a vaccine against tuberculosis, achieved by
PT reducing or eliminating expression of the alpha-crystallin heat shock
PT protein gene.
XX
PS Example; Page 25; 33pp; English.
XX
CC The invention relates to a vaccine for protection against tuberculosis in
CC which Mycobacterium sp. are attenuated by having the expression of the
CC alpha-crystallin heat shock protein gene (acr gene) reduced by at least
CC 75 percent. The attenuated strains can be used as models for screening
CC for novel antimycobacterial agents. Also, the mixture of Mycobacterium
CC sp. cellular debris can be used for detection of exposure to M. bovis and
CC M. tuberculosis, used for commercial adjuvants, used as diagnostics
CC including use as reagents for the development of monoclonal antibodies to
CC recognize wild type mycobacteria in patient samples. The vaccine is based
CC on M. tuberculosis and is obtained without the multitude of passages so

CC The present sequence represents a low oxygen induced antigen, which is
 CC used in the exemplification of the present invention.

XX Sequence 144 AA;

Query Match 100.0%; Score 736; DB 8; Length 144;

Best Local Similarity 100.0%; Pred. No. 8,1e-74; Mismatches 0; Indels 0; Gaps 0;

Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATTLPVORHRSLSLPFSESELPFAAPSPAGLRPTFDTRLMLBDEMKGRYEVRAELPGV 60
 DB 1 MATTLPVORHRSLSLPFSESELPFAAPSPAGLRPTFDTRLMLBDEMKGRYEVRAELPGV 60
 QY 61 DDDKVDIMVRDGLTTIKARTEQKDFGRSEFAYGSFVRTVSLPVGADEDDIKATYDKG 120
 DB 61 DDDKVDIMVRDGLTTIKARTEQKDFGRSEFAYGSFVRTVSLPVGADEDDIKATYDKG 120
 QY 121 ILTVSVANSEKPTKEKHQIRSTN 144
 DB 121 ILTVSVANSEKPTKEKHQIRSTN 144

RESULT 4

AAU74589 ID AAU74589 standard; protein; 330 AA.

XX AC AAU74589;

XX DT 29-AUG-2003 (revised)
 XX DT 08-MAY-2002 (first entry)

XX DE Antigenic fusion protein Erd14-DPV-MTI.

XX KW Fusion protein; tuberculosis; Mycobacterium tuberculosis;
 XX KW tuberculostatic; immunogen; vaccine; Erd14-DPV-MTI; Erd14; DPV; MTI.

XX OS Mycobacterium tuberculosis.

XX OS Chimeric.

XX PN US200209459-A1.

XX PD 24-JAN-2002.

XX PF 07-APR-1999; 99US-00287849.

XX PR 13-MAR-1997; 97US-00818112.

XX PR 01-OCT-1997; 97US-00942578.

XX PR 18-FEB-1998; 98US-00025197.

XX PR 07-APR-1998; 98US-00056556.

XX PR 30-DEC-1998; 98US-00223040.

XX PA (REED/) REED S G.

XX PA (SKEIT/) SKEIT Y A.

XX PA (DILL/) DILLON D C.

XX PA (ALDE/) ALDERSON M.

XX PA (CAMP/) CAMPOS-NETO A.

XX PI Reed SG, Skeiky YA, Dillon DC, Alderson M, Campos-Neto A;

XX PS Claim 1; Fig 2; 62pp; English.

XX CC The invention relates to a purified polypeptide which induces an immune
 CC response of Mycobacterium tuberculosis. Polypeptides of the invention are
 CC useful for diagnosing, treating or preventing M. tuberculosis infection,
 CC particularly tuberculosis infection. In particular, the polypeptides are
 CC useful as a vaccine formulation with an adjuvant to afford long-term

CC protection in animals against the development of tuberculosis. The
 CC protein coding sequence may be used to encode a protein product for use
 CC as an immunogen to induce and/or enhance an immune response to M.
 CC tuberculosis. This sequence represents an M. tuberculosis fusion protein
 CC of the invention. (Updated on 29-AUG-2003 to standardise OS field)

XX Sequence 330 AA;

Query Match 100.0%; Score 736; DB 5; Length 330;

Best Local Similarity 100.0%; Pred. No. 2,7e-73; Mismatches 0; Indels 0; Gaps 0;

Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATTLPVORHRSLSLPFSESELPFAAPSPAGLRPTFDTRLMLBDEMKGRYEVRAELPGV 60
 DB 8 MATTLPVORHRSLSLPFSESELPFAAPSPAGLRPTFDTRLMLBDEMKGRYEVRAELPGV 67
 QY 61 DDDKVDIMVRDGLTTIKARTEQKDFGRSEFAYGSFVRTVSLPVGADEDDIKATYDKG 120
 DB 61 DDDKVDIMVRDGLTTIKARTEQKDFGRSEFAYGSFVRTVSLPVGADEDDIKATYDKG 127
 QY 121 ILTVSVANSEKPTKEKHQIRSTN 144
 DB 128 ILTVSVANSEKPTKEKHQIRSTN 151

RESULT 5

AAV32060 ID AAV32060 standard; protein; 331 AA.

XX AC AAV32060;

XX DT 17-JAN-2000 (first entry)

XX DE Mycobacterium tuberculosis antigen fusion protein Mtb39A.

XX KW Tuberculosis; antigen; fusion protein; Mtb39A; ERD14; DPV; MTI;

XX KW diagnosis; therapy; vaccine; immunogen.

XX OS Mycobacterium tuberculosis.

XX OS Mycobacterium tuberculosis.

XX PH Key Location/Qualifiers

XX FT Peptide 1..7 /note= "Met/His tag"

XX FT Protein 8..151 /note= "ERD14"

XX FT Protein 154..235 /note= "DPV"

XX FT Protein 238..331 /note= "MTI"

XX PN WO9951748-A2.

XX PD 14-OCT-1999.

XX PF 07-APR-1999; 99WO-US007717.

XX PR 07-APR-1998; 98US-00056556.

XX PR 30-DEC-1998; 98US-00223040.

XX PA (CORI-) CORIYA CORP.

XX PI Skeiky YAW, Alderson M, Campos-Neto A;

XX DR WI; 1999-601610/51.

XX DR N-PSDB; AAZ20195.

XX PT New fusion proteins useful for diagnosis, prevention and treatment of
 XX PT tuberculosis.

XX PS Claim 1; Fig 2; 83pp; English.

XX CC This sequence represents a recombinant Mycobacterium tuberculosis tri-
 CC antigen fusion protein, termed Mtb39A, composed of the antigens ERD14,

CC DPV and MTI. The fusion protein is expressed in host cells using a vector
CC carrying a polynucleotide (see AAZ20195) comprising the 3 coding
CC sequences for the antigens. The invention provides fusion proteins (see
CC AA32059-71) containing at least 2 M. tuberculosis antigens. The new
CC fusion proteins and polynucleotides encoding them are useful as vaccines
CC for preventing tuberculosis (claimed), for diagnosis (via in vitro assays
CC or intradermal skin tests for detection of anti-M. tuberculosis
CC antibodies), monitoring of disease progression, and treatment of
CC tuberculosis. They are more effective immunogens than mixtures of the
CC individual protein components

CC
XX
SQ Sequence 331 AA;

Query Match 100.0%; Score 736; DB 2; Length 331;
Best Local Similarity 100.0%; Pred. No. 2,7e-73;
Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATTLPVQRHPRSLFPFSESLFAAFSPAGLRPTFOTRLRLDEMKRGYEVRAELPGV 60
DB 8 MATTLPVQRHPRSLFPFSESLFAAFSPAGLRPTFOTRLRLDEMKRGYEVRAELPGV 67

QY 61 DPKDQVIMWRDGLTIKAERTEQKDFGRSEFAYGSFVRTVSLPVGADEDDIKATYDKG 120
DB 68 DPKDQVIMWRDGLTIKAERTEQKDFGRSEFAYGSFVRTVSLPVGADEDDIKATYDKG 127

QY 121 ILTVSAVSESGKPTKHIQIRSTN 144
DB 128 ILTVSAVSESGKPTKHIQIRSTN 151

RESULT 6
ADP69755
ID ADP69755 standard; protein, 331 AA.
XX
AC ADP69755;
XX
DT 12-FEB-2004 (first entry)
XX
DE Fusion protein #2 of M. tuberculosis antigen.
XX
KM Fusion protein; Mycobacterium tuberculosis antigen;
KM tuberculosis infection; immune response; tuberculostatic; mutant; munein.
XX
OS Synthetic.
OS Mycobacterium tuberculosis.
XX
PN US2003147911-A1.
XX
PD 07-AUG-2003.
XX
PF 05-FEB-2003; 2003US-00359460.
XX
PR 13-MAR-1997; 97US-00818112.
PR 01-OCT-1997; 97US-00942578.
PR 18-FEB-1998; 98US-00025197.
PR 07-APR-1998; 98US-00056556.
PR 30-DEC-1998; 98US-00223040.
PR 07-APR-1999; 99US-00287849.
XX
PA (CORI-) CORIXA CORP.
XX
PI Reed SG, Skeiky VA, Dillon DC, Alderson M, Campos-Neto A;
XX
DR N-PSDB; ADP69754.
XX
XX WPI; 2003-897524/82.
XX
PT New fusion proteins of Mycobacterium tuberculosis for diagnosing,
PT preventing or treating tuberculosis infection or in enhancing immune
PT responses in M. tuberculosis.
XX
XX Claim 1; SEQ ID NO 4; 135pp; English.
XX
XX The present invention relates to fusion proteins of Mycobacterium

CC tuberculosis antigens, and the polynucleotide sequences encoding them.
CC The sequences of the invention are useful in a method for preventing
CC tuberculosis by administering to a subject an amount of the fusion
CC protein or the polynucleotide that encodes the fusion protein. Also
CC disclosed is a pharmaceutical composition comprising the fusion protein
CC or the polynucleotide sequence encoding it. The fusion protein induces an
CC immune response to M. tuberculosis and can be used in the diagnosis,
CC prevention, and treatment of tuberculosis infection. The present sequence
CC represents a fusion protein of a M. tuberculosis antigen.

CC
XX
SQ Sequence 331 AA;

Query Match 100.0%; Score 736; DB 7; Length 331;
Best Local Similarity 100.0%; Pred. No. 2,7e-73;
Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATTLPVQRHPRSLFPFSESLFAAFSPAGLRPTFOTRLRLDEMKRGYEVRAELPGV 60
DB 8 MATTLPVQRHPRSLFPFSESLFAAFSPAGLRPTFOTRLRLDEMKRGYEVRAELPGV 67

QY 61 DPKDQVIMWRDGLTIKAERTEQKDFGRSEFAYGSFVRTVSLPVGADEDDIKATYDKG 120
DB 68 DPKDQVIMWRDGLTIKAERTEQKDFGRSEFAYGSFVRTVSLPVGADEDDIKATYDKG 127

QY 121 ILTVSAVSESGKPTKHIQIRSTN 144
DB 128 ILTVSAVSESGKPTKHIQIRSTN 151

RESULT 7
AAV32069
ID AAV32069 standard; protein, 368 AA.
XX
AC AAV32069;
XX
DT 17-JAN-2000 (first entry)
XX
DE Mycobacterium tuberculosis antigen fusion protein Mtb36f.
XX
KM Tuberculosis; antigen; fusion protein; Mtb36f; ERD14; DPV; MTI;
KM diagnosis; therapy; vaccine; immunogen.
XX
OS Mycobacterium tuberculosis.
XX
FH Key
FH Peptide 1..8
FT /note= "Mec/Hls tag"
FT Protein 9..151
FT /note= "ERD14"
FT Protein 154..235
FT /note= "DPV"
FT Protein 238..368
FT /note= "MTI"
XX
PN MO9951748-A2.
XX
PR 14-OCT-1999.
XX
PF 07-APR-1999; 99MO-US007717.
XX
PR 07-APR-1998; 98US-00056556.
PR 30-DEC-1998; 98US-00223040.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skeiky VAM, Alderson M, Campos-Neto A;
XX
DR WPI, 1999-601610/51.
XX
XX N-PSDB; AAZ20204.
XX
PT New fusion proteins useful for diagnosis, prevention and treatment of
PT tuberculosis.
XX

PS Claim 1; Fig 11A-B; 83pp; English.

XX This sequence represents a recombinant Mycobacterium tuberculosis tri-
CC antigen fusion protein, termed Mtb36f, composed of the antigens ERD14,
CC DPV and MTI. The fusion protein is expressed in host cells using a vector
CC carrying a polynucleotide (see AA20204) comprising the coding sequences
CC for the 3 antigens. The invention provides fusion proteins (see AA132059-
CC 71) containing at least 2 M. tuberculosis antigens. The new fusion
CC proteins and polynucleotides encoding them are useful as vaccines for
CC preventing tuberculosis (claimed), for diagnosis (via in vitro assays or
CC intradermal skin tests for detection of anti-M. tuberculosis antibodies),
CC monitoring of disease progression, and treatment of tuberculosis. They
CC are more effective immunogens than mixtures of the individual protein
CC components

XX Sequence 368 AA;

Query Match 100.0%; Score 736; DB 2; Length 368;
Best Local Similarity 100.0%; Pred. No. 3.1e-73;
Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATTLPVQRHRSLSFPEPSELFAAPSPAGLRPTFDTRLMLBDEMKGREYVRAELPGV 60
DB 9 MATTLPVQRHRSLSFPEPSELFAAPSPAGLRPTFDTRLMLBDEMKGREYVRAELPGV 68
QY 61 DPDKVDIMVWDGQLTIKAERTEQKDPGRSEFAYGSFVRTVSLPVGADEDDIKATYDKG 120
DB 69 DPDKVDIMVWDGQLTIKAERTEQKDPGRSEFAYGSFVRTVSLPVGADEDDIKATYDKG 128
QY 121 ILTVSAVSESGKPTKEKHIIQIRSTN 144
DB 129 ILTVSAVSESGKPTKEKHIIQIRSTN 152

RESULT 8
AAU74598
ID AAU74598 standard; protein; 368 AA.

XX AAU74598;

XX 29-AUG-2003 (revised)
DT 08-MAY-2002 (first entry)

XX Antigenic fusion protein Erd14-DPV-MTI (Mtb36f).

XX Fusion protein; tuberculosis; Mycobacterium tuberculosis;
KM tuberculostatic; immunogen; vaccine; Erd14-DPV-MTI; Mtb36f.

XX Mycobacterium tuberculosis.
OS Chimeric.

XX US2002009459-A1.

XX 24-JAN-2002.

XX 07-APR-1999; 99US-00287849.

XX 13-MAR-1997; 97US-00818112.

XX 01-OCT-1997; 97US-00942578.

XX 18-FEB-1998; 98US-00025197.

XX 07-APR-1998; 98US-00056556.

XX 30-DEC-1998; 98US-00223040.

XX (REED/) REED S G.
PA (SKEI/) SKEIKY Y A.
PA (DILL/) DILLON D C.
PA (ALDE/) ALDERSON M.
PA (CAMP/) CAMPOS-NETO A.

PI Reed SG, Skeiky YA, Dillon DC, Alderson M, Campos-Neto A;

XX WPI; 2002-171134/22.

DR N-PSDB; ABK14138.

XX New fusion proteins of Mycobacterium tuberculosis antigens, useful for
PT diagnosing, treating or preventing M. tuberculosis infection,
PT particularly as vaccine for treating or preventing tuberculosis.

PS Claim 1; Fig 11; 62pp; English.

XX The invention relates to a purified polypeptide which induces an immune
CC response of Mycobacterium tuberculosis. Polypeptides of the invention are
CC useful for diagnosing, treating or preventing M. tuberculosis infection,
CC particularly tuberculosis infection. In particular, the polypeptides are
CC useful as a vaccine formulation with an adjuvant to afford long-term
CC protection in animals against the development of tuberculosis. The
CC protein coding sequence may be used to encode a protein product for use
CC as an immunogen to induce and/or enhance an immune response to M.
CC tuberculosis. This sequence represents an M. tuberculosis fusion protein
CC of the invention. (Updated on 29-AUG-2003 to standardise OS field)

XX Sequence 368 AA;

Query Match 100.0%; Score 736; DB 5; Length 368;
Best Local Similarity 100.0%; Pred. No. 3.1e-73;
Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATTLPVQRHRSLSFPEPSELFAAPSPAGLRPTFDTRLMLBDEMKGREYVRAELPGV 60
DB 9 MATTLPVQRHRSLSFPEPSELFAAPSPAGLRPTFDTRLMLBDEMKGREYVRAELPGV 68
QY 61 DPDKVDIMVWDGQLTIKAERTEQKDPGRSEFAYGSFVRTVSLPVGADEDDIKATYDKG 120
DB 69 DPDKVDIMVWDGQLTIKAERTEQKDPGRSEFAYGSFVRTVSLPVGADEDDIKATYDKG 128
QY 121 ILTVSAVSESGKPTKEKHIIQIRSTN 144
DB 129 ILTVSAVSESGKPTKEKHIIQIRSTN 152

RESULT 9
ADF69775
ID ADF69775 standard; protein; 368 AA.

XX ADF69775;

XX 12-FEB-2004 (first entry)

XX Fusion protein #11 of M. tuberculosis antigen.

XX Fusion protein; Mycobacterium tuberculosis antigen;
KM tuberculosis infection; immune response; tuberculostatic; mutant; mutein.

XX Synthetic.

XX Mycobacterium tuberculosis.

XX US2003147911-A1.

XX 07-AUG-2003.

XX 05-FEB-2003; 2003US-00359460.

XX 13-MAR-1997; 97US-00818112.

XX 01-OCT-1997; 97US-00942578.

XX 18-FEB-1998; 98US-00025197.

XX 07-APR-1998; 98US-00056556.

XX 30-DEC-1998; 98US-00223040.

XX 07-APR-1999; 99US-00287849.

XX (CORI-) CORIXA CORP.

PI Reed SG, Skeiky YA, Dillon DC, Alderson M, Campos-Neto A;

XX WPI; 2003-897524/82.

DR N-PSDB; ADF69774.

PT New fusion proteins of Mycobacterium tuberculosis for diagnosing,
PT preventing or treating tuberculosis infection or in enhancing immune
PT responses in M. tuberculosis.
XX
PS Claim 1; SEQ ID NO 24; 135pp; English.
XX
CC The present invention relates to fusion proteins of Mycobacterium
CC tuberculosis antigens, and the polynucleotide sequences encoding them.
CC The sequences of the invention are useful in a method for preventing
CC tuberculosis by administering to a subject an amount of the fusion
CC protein or the polynucleotide that encodes the fusion protein. Also
CC disclosed is a pharmaceutical composition comprising the fusion protein
CC or the polynucleotide sequence encoding it. The fusion protein induces an
CC immune response to M. tuberculosis and can be used in the diagnosis,
CC prevention, and treatment of tuberculosis infection. The present sequence
CC represents a fusion protein of a M. tuberculosis antigen.
XX
SQ Sequence 368 AA;

Query Match 100.0%; Score 736; DB 7; Length 368;
Best Local Similarity 100.0%; Pred. No. 3.1e-73;
Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATTLPVQRHPRSLPFPFSELFAAPSPFAGLRPTFTRLMLRLEDEMKEGRYEVRALPGV 60
Db 9 MATTLPVQRHPRSLPFPFSELFAAPSPFAGLRPTFTRLMLRLEDEMKEGRYEVRALPGV 68

QY 61 DPDKVDIMVRDGLTIKARTEKQDPGRSEFAYGSFVRTVSLPVGADDDIKATYDKG 120
Db 69 DPDKVDIMVRDGLTIKARTEKQDPGRSEFAYGSFVRTVSLPVGADDDIKATYDKG 128

QY 121 ILTVSAVSEKGPTEKHIQIRSTN 144
Db 129 ILTVSAVSEKGPTEKHIQIRSTN 152

RESULT 10
AAV32065
ID AAV32065 standard; protein; 433 AA.
XX
AC AAV32065;
XX
DT 17-JAN-2000 (first entry)
XX
DE Mycobacterium tuberculosis antigen fusion protein Mtb46f.
XX
KW Tuberculosis; antigen; fusion protein; Mtb46f; ERD14; DPV; MTI; MSI;
KW diagnosis; therapy; vaccine; immunogen.
XX
OS Mycobacterium tuberculosis.
XX
FH Key Location/Qualifiers
FT 1..8
FT Peptide /note= "Met/His tag"
FT Protein /note= "ERD14"
FT Protein /note= "DPV"
FT Protein /note= "MTI"
FT Protein /note= "MSI"
FT Protein /note= "MSL"
XX
PN MO9951748-A2.
XX
PD 14-OCT-1999.
XX
PF 07-APR-1999; 99WO-US007717.
XX
PR 07-APR-1998; 98US-00056556.
PR 30-DEC-1998; 98US-00223040.
XX
PA (CORI-) CORIXA CORP.

XX
XX Skeiky YAM, Alderson M, Campos-Neco A;
PI WPI; 1999-601610/51.
DR N-PSDB; AAZ20200.
XX
XX New fusion proteins useful for diagnosis, prevention and treatment of
PT tuberculosis.
PT
PS Claim 1; Fig 7A-B; 83pp; English.
XX
CC This sequence represents a recombinant Mycobacterium tuberculosis tetra-
CC antigen fusion protein, termed Mtb46f, composed of the antigens ERD14,
CC DPV, MTI and MSI. The fusion protein is expressed in host cells using a
CC vector carrying a polynucleotide (see AAZ20200) comprising the coding
CC sequences for the 4 antigens. The invention provides fusion proteins (see
CC AAV32059-71) containing at least 2 M. tuberculosis antigens. The new
CC fusion proteins and polynucleotides encoding them are useful as vaccines
CC for preventing tuberculosis (claimed), for diagnosis (via in vitro assays
CC or intradermal skin tests for detection of anti-M. tuberculosis
CC antibodies), monitoring of disease progression, and treatment of
CC tuberculosis. They are more effective immunogens than mixtures of the
CC individual protein components
XX
SQ Sequence 433 AA;

Query Match 100.0%; Score 736; DB 2; Length 433;
Best Local Similarity 100.0%; Pred. No. 3.9e-73;
Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATTLPVQRHPRSLPFPFSELFAAPSPFAGLRPTFTRLMLRLEDEMKEGRYEVRALPGV 60
Db 9 MATTLPVQRHPRSLPFPFSELFAAPSPFAGLRPTFTRLMLRLEDEMKEGRYEVRALPGV 68

QY 61 DPDKVDIMVRDGLTIKARTEKQDPGRSEFAYGSFVRTVSLPVGADDDIKATYDKG 120
Db 69 DPDKVDIMVRDGLTIKARTEKQDPGRSEFAYGSFVRTVSLPVGADDDIKATYDKG 128

QY 121 ILTVSAVSEKGPTEKHIQIRSTN 144
Db 129 ILTVSAVSEKGPTEKHIQIRSTN 152

RESULT 11
AAU74594
ID AAU74594 standard; protein; 433 AA.
XX
AC AAU74594;
XX
DT 29-AUG-2003 (revised)
DT 08-MAY-2002 (first entry)
XX
DE Antigenic fusion protein Erd14-DPV-MTI-MSL (Mtb46f).
XX
KW Fusion protein; tuberculosis; Mycobacterium tuberculosis;
KW tuberculostatic; immunogen; vaccine; Erd14-DPV-MTI-MSL; Mtb46f.
XX
OS Mycobacterium tuberculosis.
OS Chimeric.
XX
PN US2002009459-A1.
XX
PD 24-JAN-2002.
XX
PF 07-APR-1999; 99US-00287849.
XX
PR 13-MAR-1997; 97US-00818112.
PR 01-OCT-1997; 97US-00942578.
PR 18-FEB-1998; 98US-00025197.
PR 07-APR-1998; 98US-00056556.
PR 30-DEC-1998; 98US-00223040.
XX
PA (REED/) REED S G.

PA (SKEI/) SKEIK Y A.
PA (DILL/) DILLON D C.
PA (ALDE/) ALDERSON M.
PA (CAMP/) CAMPOS-NETO A.
PI Reed SG, Skeiky YA, Dillon DC, Alderson M, Campos-Neto A;
XX WPI; 2002-171134/22.
DR N-PSDB; ABK14134.
XX
XX
PT New fusion proteins of Mycobacterium tuberculosis antigens, useful for
PT diagnosing, treating or preventing M. tuberculosis infection,
PT particularly as vaccine for treating or preventing tuberculosis.
XX
XX
PS Claim 1; Fig 7; 62pp; English.
XX
CC The invention relates to a purified polypeptide which induces an immune
CC response of Mycobacterium tuberculosis. Polypeptides of the invention are
CC useful for diagnosing, treating or preventing M. tuberculosis infection,
CC particularly tuberculosis infection. In particular, the polypeptides are
CC useful as a vaccine formulation with an adjuvant to afford long-term
CC protection in animals against the development of tuberculosis. The
CC protein coding sequence may be used to encode a protein product for use
CC as an immunogen to induce and/or enhance an immune response to M.
CC tuberculosis. This sequence represents an M. tuberculosis fusion protein
CC of the invention. (Updated on 29-AUG-2003 to standardise OS field)
XX
SQ Sequence 433 AA;

Query Match 100.0%; Score 736; DB 5; Length 433;
Best Local Similarity 100.0%; Pred. No. 3.9e-73;
Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MATLTPVQRHRSLEPFSESLFAAFPSFAGLRPTFDTRLMLRLEDEMEKRGYEVRAELPGV 60
DB 9 MATLTPVQRHRSLEPFSESLFAAFPSFAGLRPTFDTRLMLRLEDEMEKRGYEVRAELPGV 68
QY 61 DPDKVDIMVRDGLTIKARTEQKDFGRSEFAVGSFVRTVSLPVGADDDDIKATYDKG 120
DB 69 DPDKVDIMVRDGLTIKARTEQKDFGRSEFAVGSFVRTVSLPVGADDDDIKATYDKG 128
QY 121 ILTVSAVSESGKPTKHIQIRSTN 144
DB 129 ILTVSAVSESGKPTKHIQIRSTN 152

RESULT 12
ADF69765
ID ADF69765 standard; protein; 433 AA.
XX
AC ADF69765;
XX
XX 12-FEB-2004 (first entry)
XX
DE Fusion protein #7 of M. tuberculosis antigen.
XX
KM Fusion protein; Mycobacterium tuberculosis antigen;
KM tuberculosis infection; immune response; tuberculostatic; mutant; mutein.
XX
XX Synthetic.
OS Mycobacterium tuberculosis.
XX
PN US2003147911-A1.
XX
PD 07-AUG-2003.
XX
XX 05-FEB-2003; 2003US-00359460.
XX
XX 13-MAR-1997; 97US-00818112.
PR 01-OCT-1997; 97US-00942578.
PR 18-FEB-1998; 98US-00025197.
PR 07-APR-1998; 98US-00056556.
PR 30-DEC-1998; 98US-00223040.

PR 07-APR-1999; 99US-00287849.
XX
XX (CORI-) CORIXA CORP.
XX
XX
PI Reed SG, Skeiky YA, Dillon DC, Alderson M, Campos-Neto A;
XX WPI; 2003-897524/82.
DR N-PSDB; ADF69764.
XX
XX
XX
PT New fusion proteins of Mycobacterium tuberculosis for diagnosing,
PT preventing or treating tuberculosis infection or in enhancing immune
PT responses in M. tuberculosis.
XX
XX
PS Claim 1; SEQ ID NO 14; 135pp; English.
XX
XX
CC The present invention relates to fusion proteins of Mycobacterium
CC tuberculosis antigens, and the polynucleotide sequences encoding them.
CC The sequences of the invention are useful in a method for preventing
CC tuberculosis by administering to a subject an amount of the fusion
CC protein or the polynucleotide that encodes the fusion protein. Also
CC disclosed is a pharmaceutical composition comprising the fusion protein
CC or the polynucleotide sequence encoding it. The fusion protein induces an
CC immune response to M. tuberculosis and can be used in the diagnosis,
CC prevention, and treatment of tuberculosis infection. The present sequence
CC represents a fusion protein of a M. tuberculosis antigen.
XX
SQ Sequence 433 AA;

Query Match 100.0%; Score 736; DB 7; Length 433;
Best Local Similarity 100.0%; Pred. No. 3.9e-73;
Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MATLTPVQRHRSLEPFSESLFAAFPSFAGLRPTFDTRLMLRLEDEMEKRGYEVRAELPGV 60
DB 9 MATLTPVQRHRSLEPFSESLFAAFPSFAGLRPTFDTRLMLRLEDEMEKRGYEVRAELPGV 68
QY 61 DPDKVDIMVRDGLTIKARTEQKDFGRSEFAVGSFVRTVSLPVGADDDDIKATYDKG 120
DB 69 DPDKVDIMVRDGLTIKARTEQKDFGRSEFAVGSFVRTVSLPVGADDDDIKATYDKG 128
QY 121 ILTVSAVSESGKPTKHIQIRSTN 144
DB 129 ILTVSAVSESGKPTKHIQIRSTN 152

RESULT 13
AAV32064
ID AAV32064 standard; protein; 856 AA.
XX
AC AAV32064;
XX
XX 17-JAN-2000 (first entry)
XX
DE Mycobacterium tuberculosis antigen fusion protein Mtb88f.
XX
XX Tuberculosis; antigen; fusion protein; Mtb88f; ERD14; DPV; MTI; MSL;
KM mTCC2; diagnosis; therapy; vaccine; immunogen.
XX
XX Mycobacterium tuberculosis.
OS
XX
FH Key
FH Peptide
FT /note= "Met/His tag"
FT /note= "ERD14"
FT /note= "DPV"
FT /note= "MTI"
FT /note= "MSL"
FT /note= "mTCC2"
FT Protein

```
XX XX MO9951748-A2.
XX PD 14-OCT-1999.
XX PF 07-APR-1999; 99MO-US007717.
XX PR 07-APR-1998; 98US-00065556.
XX PR 30-DEC-1998; 98US-00223040.
XX PA (CORI-) CORIXA CORP.
XX PI Skeiky YAM, Alderson M, Campos-Neto A;
XX DR WPI; 1999-601610/51.
XX DR N-PSDB; AAZ20199.
XX PT New fusion proteins useful for diagnosis, prevention and treatment of
XX PT tuberculosis.
XX PS Claim 1; Fig 6A-B; 83pp; English.
XX PS
XX CC This sequence represents a recombinant Mycobacterium tuberculosis penta-
XX CC antigen fusion protein, termed Mtb88f, composed of the antigens Erd14,
XX CC DpV, MTI, MSU and mTCC2. The fusion protein is expressed in host cells
XX CC using a vector carrying a polynucleotide (see AAZ20199) comprising the
XX CC coding sequences for the 5 antigens. The invention provides fusion
XX CC proteins (see AAJ32059-71) containing at least 2 M. tuberculosis
XX CC antigens. The new fusion proteins and polynucleotides encoding them are
XX CC useful as vaccines for preventing tuberculosis (claimed), for diagnosis
XX CC (via in vitro assays or intradermal skin tests for detection of anti-M.
XX CC tuberculosis antibodies), monitoring of disease progression, and
XX CC treatment of tuberculosis. They are more effective immunogens than
XX CC mixtures of the individual protein components
XX SQ Sequence 856 AA;

Query Match 100.0%; Score 736; DB 2; Length 856;
Best Local Similarity 100.0%; Pred. No. 1e-72;
Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATTLPVQRHPRSLPFPESELFAAPSPFAGLRPTPTDRLMLRLEDMKEGRYEVRALPGV 60
DB 9 MATTLPVQRHPRSLPFPESELFAAPSPFAGLRPTPTDRLMLRLEDMKEGRYEVRALPGV 68
QY 61 DPKDQVDIMWRDQGLTIKARTKQKDFGRSEFAYGSFVRTVSLPVGADDDIKATYDKG 120
DB 69 DPKDQVDIMWRDQGLTIKARTKQKDFGRSEFAYGSFVRTVSLPVGADDDIKATYDKG 128
QY 121 ILTVSAVSESGKPTKEKHIIQIRSTN 144
DB 129 ILTVSAVSESGKPTKEKHIIQIRSTN 152

RESULT 14
ADF69763
ID ADF69763 standard; protein; 856 AA.
XX
XX ADF69763;
AC
XX 12-FEB-2004 (first entry)
DT
XX Fusion protein #6 of M. tuberculosis antigen.
DE
XX Fusion protein; Mycobacterium tuberculosis antigen;
KM tuberculosis infection; immune response; tuberculostatic; mutant; mutein.
XX
XX Synthetic.
OS
XX Mycobacterium tuberculosis.
XX
XX US2003147911-A1.
XX
XX 07-AUG-2003.
PD
```

```
XX XX 05-FEB-2003; 2003US-00359460.
XX PF 13-MAR-1997; 97US-00818112.
XX PR 01-OCT-1997; 97US-00942578.
XX PR 18-FEB-1998; 98US-00025197.
XX PR 07-APR-1998; 98US-00065556.
XX PR 30-DEC-1998; 98US-00223040.
XX PR 07-APR-1999; 99US-00287849.
XX PA (CORI-) CORIXA CORP.
XX PI Reed SG, Skeiky YA, Dillon DC, Alderson M, Campos-Neto A;
XX DR WPI; 2003-897524/82.
XX DR N-PSDB; ADF69762.
XX PT New fusion proteins of Mycobacterium tuberculosis for diagnosing,
XX PT preventing or treating tuberculosis infection or in enhancing immune
XX PT responses in M. tuberculosis.
XX PS Claim 1; SEQ ID NO 12; 135pp; English.
XX PS
XX CC The present invention relates to fusion proteins of Mycobacterium
XX CC tuberculosis antigens, and the polynucleotide sequences encoding them.
XX CC The sequences of the invention are useful in a method for preventing
XX CC tuberculosis by administering to a subject an amount of the fusion
XX CC protein or the polynucleotide that encodes the fusion protein. Also
XX CC disclosed is a pharmaceutical composition comprising the fusion protein
XX CC or the polynucleotide sequence encoding it. The fusion protein induces an
XX CC immune response to M. tuberculosis and can be used in the diagnosis,
XX CC prevention, and treatment of tuberculosis infection. The present sequence
XX CC represents a fusion protein of a M. tuberculosis antigen.
XX SQ Sequence 856 AA;

Query Match 100.0%; Score 736; DB 7; Length 856;
Best Local Similarity 100.0%; Pred. No. 1e-72;
Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATTLPVQRHPRSLPFPESELFAAPSPFAGLRPTPTDRLMLRLEDMKEGRYEVRALPGV 60
DB 9 MATTLPVQRHPRSLPFPESELFAAPSPFAGLRPTPTDRLMLRLEDMKEGRYEVRALPGV 68
QY 61 DPKDQVDIMWRDQGLTIKARTKQKDFGRSEFAYGSFVRTVSLPVGADDDIKATYDKG 120
DB 69 DPKDQVDIMWRDQGLTIKARTKQKDFGRSEFAYGSFVRTVSLPVGADDDIKATYDKG 128
QY 121 ILTVSAVSESGKPTKEKHIIQIRSTN 144
DB 129 ILTVSAVSESGKPTKEKHIIQIRSTN 152

RESULT 15
AAU74593
ID AAU74593 standard; protein; 859 AA.
XX
XX AAU74593;
AC
XX 29-AUG-2003 (revised)
DT
XX 08-MAY-2002 (first entry)
DT
XX Antigenic fusion protein Erd14-DpV-MTI-MSU-MTCC2 (Mcb88f).
DE
XX Fusion protein; tuberculosis; Mycobacterium tuberculosis;
KM tuberculostatic; immunogen; vaccine; Mtb88f; Erd14-DpV-MTI-MSU-MTCC2.
XX
XX Mycobacterium tuberculosis.
OS
XX Chimeric.
XX
XX Key Location/Qualifiers
FH Misc-difference 857
FT /label= OTHER
```

FT /note= "OTHER= Xaa, Xaa= In frame stop codon"

XX US2002009459-A1.

XX 24-JAN-2002.

XX 07-APR-1999; 99US-00287849.

XX 13-MAR-1997; 97US-00818112.

PR 01-OCT-1997; 97US-00942578.

PR 18-FEB-1998; 98US-00025197.

PR 07-APR-1998; 98US-00056556.

PR 30-DEC-1998; 98US-00223040.

XX (REED//) REED S G.

PA (SKEI//) SKEIKY Y A.

PA (DILL//) DILLON D C.

PA (ALDE//) ALDERSON M.

PA (CAMP//) CAMPOS-NETO A.

PI Reed SG, Skeiky YA, Dillon DC, Alderson M, Campos-Neto A;

XX WPI; 2002-171134/22.

DR N-PSDB; ABK14133.

XX New fusion proteins of Mycobacterium tuberculosis antigens, useful for

PT diagnosing, treating or preventing M. tuberculosis infection,

PT particularly as vaccine for treating or preventing tuberculosis.

XX Claim 1; Fig 6; 62pp; English.

XX The invention relates to a purified polypeptide which induces an immune

CC response of Mycobacterium tuberculosis. Polypeptides of the invention are

CC useful for diagnosing, treating or preventing M. tuberculosis infection,

CC particularly tuberculosis infection. In particular, the polypeptides are

CC useful as a vaccine formulation with an adjuvant to afford long-term

CC protection in animals against the development of tuberculosis. The

CC protein coding sequence may be used to encode a protein product for use

CC as an immunogen to induce and/or enhance an immune response to M.

CC tuberculosis. This sequence represents an M. tuberculosis fusion protein

CC of the invention. (Updated on 29-AUG-2003 to standardise OS field)

XX Sequence 859 AA;

XX Query Match 100.0%; Score 736; DB 5; Length 859;

XX Best Local Similarity 100.0%; Pred. No. 1e-72;

XX Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATTLPVQRHPRSLPFRFSELPFAFPSPAGLRPTFDTRLMLRLEDEMKRGYEVRAELPGV 60

Db 9 MATTLPVQRHPRSLPFRFSELPFAFPSPAGLRPTFDTRLMLRLEDEMKRGYEVRAELPGV 68

QY 61 DPDKVDIMVDRDGLTIKARTKQKDPGRSFPAYGSFVRTVSLPVGADEDDIKATYDKG 120

Db 69 DPDKVDIMVDRDGLTIKARTKQKDPGRSFPAYGSFVRTVSLPVGADEDDIKATYDKG 128

QY 121 ILTVSVANSEKGPTEKHQIRSTN 144

Db 129 ILTVSVANSEKGPTEKHQIRSTN 152

Search completed: March 23, 2006, 05:10:46

Job time : 40.2848 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 23, 2006, 05:11:18 ; Search time 5.72908 Seconds
(without alignments)
2418.401 Million cell updates/sec

Title: US-10-617-038-24

Perfect score: 736

Sequence: 1 MATTLPVQRHPRSLPFESE.....SVAVSEGGPKTKHIQIRSTN 144

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	736	100.0	144	2 F70942	probable hsp70 prot
2	184.5	25.1	175	2 B55337	probable hsp70 hea
3	164.5	22.4	159	2 G70939	probable hsp prote
4	156	21.2	146	2 S74956	spore protein SP21
5	155.5	21.1	143	2 T40376	heat shock protein
6	151	20.5	168	2 P95941	probable small hea
7	150.5	20.4	188	2 A49942	heat shock protein
8	148	20.1	140	2 B69496	small heat shock p
9	140.5	19.1	151	2 T05739	probable heat shoc
10	139	18.9	160	2 F82582	low molecularweig
11	137.5	18.7	142	2 T46568	small heat shock p
12	136.5	18.5	152	2 S72546	heat shock protein
13	136.5	18.5	157	2 S04939	heat shock 22K pro
14	136	18.5	159	2 S72544	heat shock protein
15	133	18.1	151	2 S21600	heat shock protein
16	133	18.1	182	2 A75436	heat shock protein
17	132	17.9	147	2 D72385	heat shock protein
18	132	17.9	214	2 S45465	heat shock protein
19	132	17.9	229	2 S58210	bradyroite-specifi
20	131.5	17.9	150	2 S72545	heat shock protein
21	131.5	17.9	160	2 T05740	heat shock protein
22	131	17.8	149	2 T04171	heat shock protein
23	130.5	17.7	159	1 CYP279	heat shock protein
24	130	17.7	151	1 HHMT17	heat shock protein
25	130	17.7	210	2 T09611	heat shock protein
26	129.5	17.6	150	2 ARI842	small heat shock p
27	128	17.4	152	2 S23212	heat shock protein
28	128	17.4	153	2 S23529	heat shock protein
29	128	17.4	156	2 S71566	heat shock protein

30	127	17.3	232	2 C96566	hypothetical prote
31	126.5	17.2	154	2 S24396	heat shock protein
32	126.5	17.2	161	2 S00646	heat shock protein
33	126	17.1	157	2 T14381	heat-shock protein
34	126	17.1	160	2 T43377	low-molecularweig
35	125.5	17.1	154	2 T44801	heat shock protein
36	125.5	17.1	155	2 T06449	probable heat shoc
37	125.5	17.1	161	2 S33566	heat shock protein
38	123.5	16.8	154	2 JS0710	heat shock protein
39	123.5	16.8	159	2 T04173	heat shock protein
40	123	16.7	157	2 S06074	heat shock protein
41	122.5	16.6	156	2 J00351	heat shock protein
42	122.5	16.6	161	2 J00352	heat shock protein
43	122.5	16.6	163	2 S71567	small heat-shock p
44	122	16.6	143	2 A57270	heat shock protein
45	121.5	16.5	150	2 S20874	heat shock protein

ALIGNMENTS

RESULT 1

probable hsp70 protein - Mycobacterium tuberculosis (strain H37Rv)
N:Alternate names: 14K antigen; 19K major membrane protein
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence revision 17-Jul-1998 #text_change 09-Jul-2004
C:Accession: F70942; A42651; A43823

R:Core, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feldwell, T.; Gentles, S.; Hamlin, N.; Holtz, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998

A:Authors: Sgares, R.; Suleston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G. A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome A:Reference number: A70500; MUID:9825987; PMID:9634230

A:Accession: F70942
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-144 <COL>

A:Cross-references: UNIPROT:P30223; UNIPARC:UP10000044BF; GB:AL021899; GB:AL123456; NID A:Experimental source: strain H37Rv
R:Verdon, A.; Hartskeerl, R.A.; Schultema, A.; Kolk, A.H.; Young, D.B.; Lachygra, R. J. Bacteriol. 174, 1352-1359, 1992

A:Title: The 14,000-molecular-weight antigen of Mycobacterium tuberculosis is related to A:Reference number: A42651; MUID:92138631; PMID:1370952

A:Accession: A42651
A:Status: preliminary
A:Molecule type: DNA

A:Residues: 1-144 <VER>
A:Cross-references: UNIPARC:UP10000044BF; GB:S79751; NID:9244561; PIDN:AAB21317.1; PID:K

A:Note: sequence extracted from NCBI backbone (NCBIN:79751, NCBI:P:79752)
R:Lee, B.Y.; Hefta, S.A.; Brennan, P.J. Infect. Immun. 60, 2066-2074, 1992

A:Title: Characterization of the major membrane protein of virulent Mycobacterium tuberc A:Reference number: A43823; MUID:9225631; PMID:1563797

A:Accession: A43823
A:Molecule type: protein
A:Residues: 2-144 <LEB>

A:Cross-references: UNIPARC:UP10000124E06
A:Experimental source: strain Erdman
C:Genetics:

A:Gene: hspX
C:Superfamily: alpha-crystallin-related small heat shock protein
C:Keywords: membrane protein; surface antigen

Query Match 100.0%; Score 736; DB 2; Length 144;
Best Local Similarity 100.0%; Pred. No. 4.8e-60;
Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	DB
1 MATTLPVQRHPRSLPFESESLFAAFSPAGLRPTFDRLRLRDEMEGRYVAAELPGV 60	1 MATTLPVQRHPRSLPFESESLFAAFSPAGLRPTFDRLRLRDEMEGRYVAAELPGV 60

Best Local Similarity 33.6%; Pred. No. 1,2e-06;
Matches 45; Conservative 26; Mismatches 47; Indels 16; Gaps 6;

Oy 5 LPVGRHPRSLPFEFSELF-----AAFPSPAGLRPTFDTRLMLRLEDMKRGYVRALP 58
 | :| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 29 LAHHENMRNLFFDPAFRSFTRLPLISGFSFGAGWPS-----VEVSDDKE--IKVTAEAPL 81
 || :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

Oy 59 GVDPDKVDIMWRDGOLTIKAERTEQKDFDSSEFAYSFPVRTLSLPVGADDEDDIKAT 116
 |:| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 82 GID-EKVVELSLSDGVLSIRGERKRAETEDQENOFSSRYGRFERRIALGYEVDESKNVAT 140
 | :| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

Oy 117 YDKGLTVSVAVSE 130
 :| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 141 FRNGVLTVTLPKTPE 154
 | :| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

RESULT 7
AAccession: A49942
heat shock protein SP21 - Stigmatella aurantiaca
N.Alternate names: spore protein
C.Species: Stigmatella aurantiaca
C.Date: 11-Nov-1994 #sequence_revision 11-Nov-1994 #text_change 09-Jul-2004
CAccession: A49942; S27669
R.Heidelberg, M.; Skladny, H.; Schaller, H.U.
U. Bacteriol. 175, 7479-7482, 1993
A.Title: Heat shock and development induce synthesis of a low-molecular-weight stress-resistance protein from *Spirillum volutans*.
A.Reference number: A49942; PMID:822695
A.Accession: A49942
A>Status: preliminary
M.Molecule type: DNA
A.Residues: 1-188 <HE>
A.Cross-references: UNIPROT:Q06823; UNIPARC:UPI0000135D57; GB:M94510; NID:g152657; PIDD:
C.Genetics:
A.Gene: hspa
C.Superfamily: alpha-crystallin-related small heat shock protein

Query Match 20.4%; Score 150.5; DB 2; Length 188;
Best Local Similarity 27.9%; Pred. No. 1.5e-06;
Matches 43; Conservative 29; Mismatches 45; Indels 37; Gaps 7;

Oy 15 PFEPSEL-----FAFPSFAGLR-----PFDTRLMKLEDEMKGREYVRALPGVD 61
 | :| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 23 FQGOMELMMWDPELLNNHFANRCPPAFVFPAFEVR-----EYEA-YIFKADLPQVD 75
 || :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

Oy 62 PDKDVIDMYRDGOLTIKAERTEQKDFDSSEFAYSFPVRTLSLPVGADDEDDIKATY 117
 ::||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 76 -EKDIENVLTIGRVSGSKRKREKRRESRFAYKXETTFGSFSAFLTPEGVDGNVAIDL 134
 || :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

Oy 118 DKGITVSY-----AVSEGKPYTEKHIIQ 139
 || :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 135 KNGVLTLLTPKKRPVOVKRIQVASSTGEQKEHIK 168
 | :| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

RESULT 8
B69436
small heat shock protein (hsp20-2) homolog - Archaeoglobus fulgidus
C.Species: Archaeoglobus fulgidus
C.Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
CAccession: B69436
R.Klien, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
.J. Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.;
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 380, 364-370, 1997
A.Author: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artach, P.; Kaine, B.P.; Sykes, S.N.;
Smith, H.O.; Woese, C.R.; Venter, J.C.
A>Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeon
A.Reference number: A69250; PMID:98049343; PMID:9389475
A.Accession: B69436
A.Status: preliminary; nucleic acid sequence not shown; translation not shown
M.Molecule type: DNA
A.Residues: 1-140 <KB>
A.Cross-references: UNIPROT:Q28308; UNIPARC:UPI0000055B10; GB:A6000967; GB:A6000782; NID:
C.Superfamily: alpha-crystallin-related small heat shock protein

```
Query Match 20.1%; Score 148; DB 2; Length 140;
Best Local Similarity 32.8%; Pred. No. 1.8e-06;
Matches 45; Conservative 20; Mismatches 58; Indels 14; Gaps 6;

Qy 9 RHRPSLPEFSEFLFAAPSPFAGLRTPTDRLRLDEMEKGRYRVARLPGVDPDKVD 67
Db 12 RRMGRFRRLLEBFRRPEVKEFRVTMPV-----DVIDEGQIRVADLPFSKE-DLE 64

Qy 68 IMVDSGQLTIKARTQ-KDPDG--RSEFVAGSFVRTVSLPVGADEDDIKATYDKGILT 123
Db 65 IYFDEGDLVITAEKKEEFKKGEVLRERBRMGKYRRIRALPAGLDDAVAKXNNGVL- 123

Qy 124 VSAVSESGPTKEKHIOI 140
Db 124 -EITIPKLKDKRKAQOI 139

RESULT 9
T05739
Probable heat shock protein 17 - barley
C/Species: Hordeum vulgare (barley)
C/Date: 09-Apr-1999 #sequence_revision 09-Apr-1999 #text_change 09-Jul-2004
R/Marmitoli, N.; Pavesti, A.; di Cola, G.; Hartings, H.; Raho, G.; Conte, M.R.; Perrotta,
Genome 36, 1111-1118, 1993
A/Title: Identification, characterization and analysis of cDNA and genomic sequences end
A/Reference number: Z15446; MUID:94156165; PMID:8112573
A/Accession: T05739
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-151 <MAB>
A/Cross-references: UNIPROT:Q40056; UNIPARC:UPI00000A2508; EMBL:X64560; NID:9509175; PID
A/Experimental source: cv. Onice
C/Genetics:
C/Superfamily: alpha-crystallin-related small heat shock protein
C/Keywords: heat shock; stress-induced protein

Query Match 19.1%; Score 140.5; DB 2; Length 151;
Best Local Similarity 32.7%; Pred. No. 9.6e-06;
Matches 50; Conservative 25; Mismatches 57; Indels 21; Gaps 8;

Qy 5 LPVGRHPRSLPFSESELA-----AFPSFAGLRPTDRLRLDEMEKGRYEV--- 53
Db 1 MSIVRSNVLDLP-FADLWADPLDTPRSIFPAISGNSSETAVBERRM--DMKGRLEAVIF 57

Qy 54 RAEI.PGVDPDKVDIMVRDQO-LTIKARTKQKDPDG---RSEFVAGSFVRTVSLPVGA 108
Db 58 KADLP.GVAKK-EVKVEVDGAVLVGERTKEKDKNDKMRVERRSKGFVRPRLPDG 116

Qy 109 DEDDIKATYDKGILTVAVSE-GKPTKEKHIOI 140
Db 117 KVDEVKAGLENGVLTIVPKAEVKKPEVKAIEI 149

RESULT 10
F82582
Low molecular weight heat shock protein Xf2234 [imported] - Xylella faecidiosa (strain 9
C/Species: Xylella faecidiosa
C/Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C/Accession: F82582
R/Anonymous, The Xylella faecidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A/Title: The genome sequence of the plant pathogen Xylella faecidiosa.
A/Reference number: A82515; MUID:20365717; PMID:10910347
A/Note: for a complete list of authors see reference number A59328 below
A/Accession: F82582
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-160 <SIM>
A/Cross-references: UNIPROT:Q9PBH0; UNIPARC:UPI00000C239B; GB:AB004036; GB:AB003849; NID
A/Experimental source: strain 945c
```

```
R/Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; Al
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, U.E.A.; Carraro, D.M.; Carter, H
as-Melo, E.; Docena, C.; El-Dozry, H.; Facincan, A.P.; Ferreira, A.J.S.
Submitted to GenBank, June 2000
A/Authors: Ferreira, V.C.A.; Parro, J.A.; Fraga, J.S.; Franco, S.C.; Franco, M.C.; Frohme
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitchina, J.P.; Krieger, J.E.; Kurame, E.E.; Laigre
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Martino, C.L.; Marques, M.V.; Martins, E
A/Authors: Martins, E.M.F.; Marukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J.; de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A/Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tsubako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A/Reference number: A59328
A/Contents: annotation
C/Genetics:
A/Gene: XF2234
C/Superfamily: alpha-crystallin-related small heat shock protein

Query Match 18.9%; Score 139; DB 2; Length 160;
Best Local Similarity 32.7%; Pred. No. 1.4e-05;
Matches 34; Conservative 24; Mismatches 40; Indels 6; Gaps 3;

Qy 41 RLDEMEKGRYEVRAELPGVDPDKVDIMVRDQLTIKAR---TEQKDPDGRSEPAYG 96
Db 46 RVDIKERPNQVLYADLPGLDP-ADIEVQMDKGLTSIKGRKTESSQTEHFRIRERYG 104

Qy 97 SFVRTVSLPVGADEDDIKATYDKGILTVAVSESGPTKEKHIOI 140
Db 105 SFRRFALPDSADADGITAGSHGVLSIFIP-KRAATTPRIQV 147

RESULT 11
T46658
small heat shock protein [validated] - Thermotoga maritima (DSM 3109)
C/Species: Thermotoga maritima
A/Variety: DSM 3109
C/Date: 18-Feb-2000 #sequence_revision 18-Feb-2000 #text_change 09-Jul-2004
C/Accession: T46658
R/Michellin, E.T.; Flynn, G.C.
Submitted to the EMBL Data Library, November 1998
A/Description: The unique chaperone operon of Thermotoga maritima: cloning and initial c
A/Reference number: Z23119
A/Accession: T46658
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-142 <MIC>
A/Cross-references: UNIPROT:Q92FD1; UNIPARC:UPI00000BD94; EMBL:AF106330; PIDN:AACT9726.1
A/Experimental source: DSM 3109
C/Genetics:
A/Gene: shsp
C/Function:
A/Description: molecular chaperone [validated, MUID:99328968]
C/Superfamily: alpha-crystallin-related small heat shock protein

Query Match 18.7%; Score 137.5; DB 2; Length 142;
Best Local Similarity 29.7%; Pred. No. 1.7e-05;
Matches 41; Conservative 25; Mismatches 53; Indels 19; Gaps 5;

Qy 7 VQHRPSLPEFSEFLFAAPSPFAGLRTPTDRLRLDEMEKGRYEVRAELPGVDPDKVD 66
Db 16 LQREIDLPLDFDR-----TEVRAPRMDVPERIDDEV-----VIEVILPGID-RKDV 61

Qy 67 DIMVDSGQLTIKARTQKDPDGRS---EFVAGSFVRTVSLPVGADEDDIKATYDKGILT 122
Db 62 QITVEENILKISEKLEKRGKNGYVYVSAGKFERAIRLPDYDVVERIKAEYKNGVL 121

Qy 123 TVSAVSESGPTKEKHIOI 140
Db 122 TIRVPKKEER-KKKVIEV 138

RESULT 12
S72546
```

heat shock protein 17.0 - pearl millet

C:Species: Pennisetum americanum (pearl millet)

C>Date: 14-Apr-1998 #sequence_revision 15-May-1998 #text_change 09-Jul-2004

C:Accession: S72546

R:Cavan, G.P.; Skoet, K.; Stevens, M.J.; Howarth, C.J.

A:Title: Sequence announcement.

A:Reference number: S72545

A:Accession: S72546

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: mRNA

A:Residues: 1-152 <CAV>

A:Cross-references: UNIPROT:Q40866; UNIPARC:UPI000009DD14; EMBL:X94191; NID:g1122314; PI

A>Note: the nucleotide sequence was submitted to the EMBL Data Library, December 1995

A>Note: the source is designated as Pennisetum glaucum

C:Superfamily: alpha-crystallin-related small heat shock protein

Query Match 18.5%; Score 136.5; DB 2; Length 152;

Best Local Similarity 32.4%; Pred. No. 2.2e-05;

Matches 44; Conservative 24; Mismatches 57; Indels 11; Gaps 6;

14 LFPFSELPFAFPSPAGLRPTFDRLM---RLDEMKGRYEVRAELPGVDPKVDIMV 70

Db 17 LWDPPDSMFRSIVOSAG-SPSDTAFAAARIDMKETPEAHVFKADLPVKKE-EVKVEV 74

Qy 71 RDGQ-LTIKARTEOKDFG---RSEFAYGSFVTVSLPVGADDDIKATYDKGILTVS 125

Db 75 EDGNVLTVISGQSKSKEDKNDRMKRVKRSRSGQFMKRPFLPGNAKVDQYKAGLENGVLTVT 134

Qy 126 V-AVSEGGKPTKHIQI 140

Db 135 VPKAEEKKPEVKAI 150

Db 135 VPKAEEKKPEVKAI 150

RESULT 13

heat shock 22K protein - Chlamydomonas reinhardtii

C:Species: Chlamydomonas reinhardtii

C>Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 09-Jul-2004

C:Accession: S04939

R:Griffith, B.; Ish-Shalom, D.; Even, D.; Glaczinski, H.; Ottersbach, P.; Ohad, I.; Kloppe

Eur. J. Biochem. 182, 539-546, 1989

A:Title: The nuclear-coded chloroplast 22-kDa heat-shock protein of Chlamydomonas. Evid

A:Reference number: S04939; MUID:89325298; PMID:2473899

A:Accession: S04939

A:Molecule type: mRNA

A:Residues: 1-157 <GRI>

A:Cross-references: UNIPROT:P12811; UNIPARC:UPI000012CC15; EMBL:X15053; NID:g18151; PIDN

C:Superfamily: alpha-crystallin-related small heat shock protein

C:Keywords: chloroplast

Query Match 18.5%; Score 136.5; DB 2; Length 157;

Best Local Similarity 38.8%; Pred. No. 2.3e-05;

Matches 38; Conservative 14; Mismatches 37; Indels 9; Gaps 4;

Qy 51 YEVAELPGVDPDDKVDIMVDRDGLTIKAEK---EQKDPDG---RSEFAYGSFVTVSL 104

Db 58 FELHADAAGMBD-DVKVLEGLVMTVGERKLSHTTEAGGKWRSEKTRVAFSL 116

Qy 105 PVGADDDIKATYDKGILTVSAVSE--GKPTKHIQI 140

Db 117 PENANPDGITAMDKGVLVTVPRKRPAPKPEKRIAV 154

Db 117 PENANPDGITAMDKGVLVTVPRKRPAPKPEKRIAV 154

RESULT 14

S72544

heat shock protein 17.9 - pearl millet

C:Species: Pennisetum americanum (pearl millet)

C>Date: 01-May-1998 #sequence_revision 01-May-1998 #text_change 09-Jul-2004

C:Accession: S72544

R:Cavan, G.P.; Skoet, K.; Stevens, M.J.; Howarth, C.J.

A:Title: Sequence announcement.

A:Reference number: S72544

A:Accession: S72544

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: mRNA

A:Residues: 1-159 <CAV>

A:Cross-references: UNIPROT:Q40867; UNIPARC:UPI000008AAF7; EMBL:X94193; NID:g1122316; PI

A>Note: the nucleotide sequence was submitted to the EMBL Data Library, December 1995

A>Note: the source is designated as Pennisetum glaucum

C:Genetics: hsp17.9

C:Superfamily: alpha-crystallin-related small heat shock protein

C:Keywords: heat shock; stress-induced protein

Query Match 18.5%; Score 136; DB 2; Length 159;

Best Local Similarity 32.8%; Pred. No. 2.6e-05;

Matches 44; Conservative 21; Mismatches 53; Indels 16; Gaps 5;

Qy 13 SLFPFSELPFAFPSPAGLRPTFDRLMRLDEMKGRYEVRAELPGVDPKVDIMVDR 72

Db 34 SLFPSPPTSETAFAG-----ARIDMKETPEAHVFKADVPALKKE-EVKVEVD 83

Qy 73 GQ-LTIKAEK---EQKDPGRSEFAYGSFVTVSLPVGADDDIKATYDKGILTVSVA 127

Db 84 GNVLIQISGERKKEDEKTDVHVRSSGKFWRRFRLPENAKTQIRASMEGVLTVTV 143

Qy 128 VSE-GKPTKHIQI 140

Db 144 KEVKKPEVKSIOI 157

Db 144 KEVKKPEVKSIOI 157

RESULT 15

heat shock protein 16.9B - wheat

C:Species: Triticum aestivum (common wheat)

C>Date: 08-Jun-1994 #sequence_revision 12-May-1995 #text_change 09-Jul-2004

C:Accession: S21600

R:Weng, J.; Wang, Z.F.; Nguyen, H.T.

submitted to the EMBL Data Library, February 1992

A:Description: Cloning and characterization of cytoplasmic LMW HSP genes from wheat.

A:Reference number: S21600

A:Accession: S21600

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-151 <MEN>

A:Cross-references: UNIPROT:Q41560; UNIPARC:UPI000009F825; EMBL:X64618; NID:g21804; PIDN

C:Superfamily: alpha-crystallin-related small heat shock protein

Query Match 18.1%; Score 133; DB 2; Length 151;

Best Local Similarity 31.5%; Pred. No. 4.6e-05;

Matches 45; Conservative 29; Mismatches 53; Indels 16; Gaps 7;

Qy 13 SLFPFSELPFAFPSPAGLRPTF-----DTRLM---RLDEMKGRYEVRAELPGVD 63

Db 8 NVFPFPLMDPDPTRSFVPAISGGSETAFAANRMWKEKTPPEAHVFKADLPVKKE 67

Qy 64 KDQVIMRDPGQ-LTIKARTEOKDFG---RSEFAYGSFVTVSLPVGADDDIKATYD 118

Db 68 -EVAVEYEDGNVLVSGERTKEKDKNDKMHVRSSGKFWRRFRLLEDKVEEVKAGLE 126

Qy 119 KGILTVSAVSE-GKPTKHIQI 140

Db 127 NGVLTVTVPKAEVKKPEVKAIQI 149

Db 127 NGVLTVTVPKAEVKKPEVKAIQI 149

Search completed: March 23, 2006, 05:26:30

Job time : 5.72908 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocollaboration Ltd.

OM protein - protein search, using SW model

Run on: March 23, 2006, 04:56:43 ; Search time 36.8238 Seconds
(without alignments)
2758.529 Million cell updates/sec

Title: US-10-617-038-24
Perfect score: 736
Sequence: 1 MATTLPVGRHRSRLFFPFSF.....SVAVSEKPTKHIQIRSTN 144

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues
Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_05.80.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	731	99.3	143	1 14KD_MYCBO	POA5b8 mycobacteri
2	731	99.3	143	1 14KD_MYCTU	POA5b7 mycobacteri
3	442	60.1	86	2 08KY85_MYCTU	08KY85 mycobacteri
4	442	60.1	86	2 08KY81_MYCS2	08KY81 mycobacteri
5	442	60.1	86	2 08KY73_MYCTT	08KY73 mycobacteri
6	442	60.1	86	2 08KY82_MYCGO	08KY82 mycobacteri
7	442	60.1	86	2 08KY70_MYCCO	08KY70 mycobacteri
8	442	60.1	86	2 08KY76_MYCGN	08KY76 mycobacteri
9	442	60.1	86	2 08KY72_MYCCN	08KY72 mycobacteri
10	441	59.9	86	2 08KY68_MYCCH	08KY68 mycobacteri
11	441	59.9	86	2 08KY75_MYCPA	08KY75 mycobacteri
12	434	59.0	86	2 08KY69_MYCFO	08KY69 mycobacteri
13	363	49.3	86	2 08KY67_MYCXE	08KY67 mycobacteri
14	343	46.6	86	2 08KY84_MYCNO	08KY84 mycobacteri
15	343	46.6	86	2 08KY71_MYCGS	08KY71 mycobacteri
16	340	46.2	86	2 08KY80_MYCNR	08KY80 mycobacteri
17	340	46.2	86	2 08KY79_MYCUL	08KY79 mycobacteri
18	336	45.7	86	2 08KY78_MYCUL	08KY78 mycobacteri
19	333	45.2	86	2 08KY77_MYCUL	08KY77 mycobacteri
20	235.5	32.0	139	2 04NAQ3_SMICC	Q4naq3 mycobacteri
21	232	31.5	157	2 04NM99_SMICC	Q4nm99 mycobacteri
22	188	25.5	48	2 08KY74_MYCHA	08KY74 mycobacteri
23	187	25.4	166	2 08N5S5_RHOXA	Q8n5s5 rhodospirill
24	184.5	25.1	175	2 092287_RHIME	Q92287 rhizobium m
25	184	25.0	48	2 08KHQ1_MYCUL	08KHQ1 mycobacteri
26	184	25.0	48	2 08KY83_MYCNO	08KY83 mycobacteri
27	173	23.5	173	2 0891A9_BRASA	Q891a9 bradyrhizob
28	171.5	23.3	147	2 074FR8_GEOSL	Q74fr8 geobacter s
29	170	23.1	146	2 073TL7_MYCPA	Q73tl7 mycobacteri
30	166.5	22.6	231	2 088HW9_PSEPK	Q88hw9 pseudomonas
31	165.5	22.5	154	2 04UKR8_RICFE	Q4ukr8 rickettsia

32	165	22.4	199	2	041YM3_AZOVI	Q41ym3 azotobacter
33	164.5	22.4	147	2	06MJP8_BDEBA	Q6mjp8 bdellovibri
34	164.5	22.4	159	2	07U2G3_MYCBO	Q7u2g3 mycobacteri
35	164.5	22.4	159	2	053673_MYCTU	Q53673 mycobacteri
36	161.5	21.9	189	2	05K552_9BACT	Q5k552 uncultured
37	161	21.9	142	2	05NE64_FRATT	Q5ne64 francisella
38	160.5	21.8	166	2	04TIT0_GSPFN	Q4tit0 erythrobact
39	158.5	21.5	158	2	04HAV5_9DBIO	Q4hav5 delnoccocus
40	158.5	21.5	188	2	052TH0_LEGPH	Q52th0 legionella
41	158	21.5	145	2	09XDG8_MYCSM	Q9xdg8 mycobacteri
42	156	21.2	146	2	F72977_SYNX3	P72977 synechocyst
43	155.5	21.1	140	2	06ACP4_LEIYX	Q6acf4 leifsonia x
44	155.5	21.1	143	1	HSP16_SCHPO	O14368 schistosach
45	155.5	21.1	151	2	Q7UF81_RHOBA	Q7uf81 rhodospirilli

ALIGNMENTS

RESULT 1
ID 14KD_MYCBO STANDARD; PRT; 143 AA.
AC POA5B8; P30223;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE 14 kDa antigen (16 kDa antigen) (HSP 16.3).
GN Name=hspX; OrderedLocustNames=Mb2057c;
OS Mycobacterium bovis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium tuberculosis complex.
OX NCBI_TaxID=1765;
RN [1]
RP NUCLEOTIDE SEQUENCE (LARGE SCALE GENOMIC DNA).
RC STRAIN=AF2122/97; PubMed=12788972; DOI=10.1073/pnas.1130426100;
RX MEDLINE=22709107; PubMed=12788972; DOI=10.1073/pnas.1130426100;
RA Garnier T., Eigmeier K., Camus J.-C., Medina N., Mansoor H.,
RA Pryor M., Duthey S., Grondin S., Lacroix C., Monempe C., Simon S.,
RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA Parkhill J., Barrall B.G., Cole S.T., Gordon S.V., Hewison R.G.;
RT "The complete genome sequence of Mycobacterium bovis";
RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882 (2003).
RL -1- SUBCELLULAR LOCATION: Probably the external side of the cell wall
(By similarity).
CC -1- SIMILARITY: Belongs to the small heat shock protein (HSP20)
family.
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL: BX248341; CAD96910.1; -; Genomic_DNA.
DR InterPro: IPR002068; HSP20.
DR Pfam: PF00011; HSP20.1.
DR PROSITE: PS01031; HSP20.1.
KM Antigen; Complete proteome.
FT INT MET 0 By similarity.
SQ SEQUENCE 143 AA; 16096 MW; 751AD94203226C9 CRC64;
Query Match 99.3%; Score 731; DB 1; Length 143;
Best Local Similarity 100.0%; Pred. No. 2e-58;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 ATTLPVGRHRSRLFFPFSFAGLRPTFDRLMKLEDMKGRYEVRAELPGVD 61
DB 1 ATTLPVGRHRSRLFFPFSFAGLRPTFDRLMKLEDMKGRYEVRAELPGVD 60
QY 62 PDKVDIMVDDGQITTAERTKQDFGRSEFAYGSVVRVLSLPVGDDEDDIATYKGI 121
DB 61 PDKVDIMVDDGQITTAERTKQDFGRSEFAYGSVVRVLSLPVGDDEDDIATYKGI 120

Qy	122	LTVSVAVSEGKPTKHIQIRSTN	144
Db	121	LTVSVAVSEGKPTKHIQIRSTN	143

RESULT 2

1_KD_MYCTU	STANDARD:	PRT:	143 AA.
ID	14_KD_MYCNU		
AC	POA5E7; P30223;		
DT	01-APR-1993 (Rel. 25, Created)		
DT	01-JUL-1993 (Rel. 26, Last sequence update)		
DT	10-MAY-2005 (Rel. 47, Last annotation update)		
DE	14 kDa antigen (16 kDa antigen) (HSP 15.3)		
GN	Name=nspX; Order=relLocusNames=rv2031c, MT2090; ORFNames=MTV018.18c		
OS	Mycobacterium tuberculosis.		
OC	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;		
OC	Corynebacteriaceae; Mycobacteriaceae; Mycobacterium;		
OC	Mycobacterium tuberculosis complex.		
NCBI_TaxId=1773;			

CC This Swiss-Pat entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.

DR EMBL, S79751; AAB21317.1; -; Genomic_DNA.
DR EMBL, M76712; AAA25342.1; -; Genomic_DNA.
DR EMBL, BX842578; CAA17245.1; -; Genomic_DNA.
DR EMBL, AE007059; AAK46369.1; -; Genomic_DNA.
DR PIR, F70942; F70942; -;
DR TIGR, MT2090; -;
DR Tubercula1c; Rv2031c; -;
DR InterPro; IPR002068; Hsp20.
DR Pfam; PF00011; HSP20.1.
DR PROSITE; PS01031; HSP20.1.
KW Antigen; Complete proteome; Direct protein sequencing
FT INIT MET 0 0
SQ SEQUENCE 143 AA; 16096 MW; 751AD94203226CE9 CRC64

RESULT 3	
Q8KY85 MYCTU	
ID Q8KY85 MYCTU PRELIMINARY;	PRT; 86 AA

QY 88 DGRSEFAYGSFVRTVSLPVGADDDI 113
61 DGRSEFAYGSFVRTVSLPVGADDDI 86

RESULT 4

ID Q8KY81_MYCSZ PRELIMINARY; PRT; 86 AA.
AC Q8KY81;
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Alpha-crystallin protein (Fragment).
GN Name=acr;
OS Mycobacterium szulgai.
OC Bacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Mycobacterium.
OX NCBI_TaxID=1787;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=NTC 10831;
RA Shampura I.C., Bastian I., Fonteyne P.-A., Portaeis F.;
RL Submitted (APR-2000) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: Belongs to the small heat shock protein (HSP20) family.
DR EMBL; AF253448; AAM69210.1; -; Genomic_DNA.
DR InterPro; IPR002068; Hsp20.
DR Pfam; PF00011; HSP20; 1.
DR PROSITE; PS01031; HSP20; 1.
FT NON_TER 1
SQ SEQUENCE 86 AA; 9814 MW; 1D6966C08DE5845B CRC64;

Query Match 60.1%; Score 442; DB 2; Length 86;
Best Local Similarity 100.0%; Pred. No. 1.9e-32;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28 FAGLRPTFDTRLMLRLEDEMKRGYEVRAELPGVDPDKVDIMVRDGLTIKAERTEQKF 87
Db 1 FAGLRPTFDTRLMLRLEDEMKRGYEVRAELPGVDPDKVDIMVRDGLTIKAERTEQKF 60

QY 88 DGRSEFAYGSFVRTVSLPVGADDDI 113
61 DGRSEFAYGSFVRTVSLPVGADDDI 86

RESULT 5

ID Q8KY73_MYCIT PRELIMINARY; PRT; 86 AA.
AC Q8KY73;
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Alpha-crystallin protein (Fragment).
GN Name=acr;
OS Mycobacterium intracellulare.
OC Bacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Mycobacterium;
OC Mycobacterium avium complex (MAC).
OX NCBI_TaxID=1767;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=M4199;
RA Shampura I.C., Bastian I., Fonteyne P.-A., Portaeis F.;
RL Submitted (APR-2000) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: Belongs to the small heat shock protein (HSP20) family.
DR EMBL; AF253459; AAM69221.1; -; Genomic_DNA.
DR InterPro; IPR002068; Hsp20.
DR Pfam; PF00011; HSP20; 1.
DR PROSITE; PS01031; HSP20; 1.
FT NON_TER 1

FT NON_TER 86
SQ SEQUENCE 86 AA; 9814 MW; 1D6966C08DE5845B CRC64;
Query Match 60.1%; Score 442; DB 2; Length 86;
Best Local Similarity 100.0%; Pred. No. 1.9e-32;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28 FAGLRPTFDTRLMLRLEDEMKRGYEVRAELPGVDPDKVDIMVRDGLTIKAERTEQKF 87
Db 1 FAGLRPTFDTRLMLRLEDEMKRGYEVRAELPGVDPDKVDIMVRDGLTIKAERTEQKF 60

QY 88 DGRSEFAYGSFVRTVSLPVGADDDI 113
61 DGRSEFAYGSFVRTVSLPVGADDDI 86

RESULT 6

ID Q8KY82_MYCGO PRELIMINARY; PRT; 86 AA.
AC Q8KY82;
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Alpha-crystallin protein (Fragment).
GN Name=acr;
OS Mycobacterium gordonae.
OC Bacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Mycobacterium.
OX NCBI_TaxID=1778;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC14470;
RA Shampura I.C., Bastian I., Fonteyne P.-A., Portaeis F.;
RL Submitted (APR-2000) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: Belongs to the small heat shock protein (HSP20) family.
DR EMBL; AF253447; AAM69209.1; -; Genomic_DNA.
DR InterPro; IPR002068; Hsp20.
DR Pfam; PF00011; HSP20; 1.
DR PROSITE; PS01031; HSP20; 1.
FT NON_TER 1
SQ SEQUENCE 86 AA; 9814 MW; 1D6966C08DE5845B CRC64;

Query Match 60.1%; Score 442; DB 2; Length 86;
Best Local Similarity 100.0%; Pred. No. 1.9e-32;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28 FAGLRPTFDTRLMLRLEDEMKRGYEVRAELPGVDPDKVDIMVRDGLTIKAERTEQKF 87
Db 1 FAGLRPTFDTRLMLRLEDEMKRGYEVRAELPGVDPDKVDIMVRDGLTIKAERTEQKF 60

QY 88 DGRSEFAYGSFVRTVSLPVGADDDI 113
61 DGRSEFAYGSFVRTVSLPVGADDDI 86

RESULT 7

ID Q8KY70_MYCO PRELIMINARY; PRT; 86 AA.
AC Q8KY70;
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Alpha-crystallin protein (Fragment).
GN Name=acr;
OS Mycobacterium lentiflavum.
OC Bacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Mycobacterium;
OX NCBI_TaxID=141349;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=M96-1701;

QY 88 DGRSEFAYGSFVRTVSLPVGADDDI 113
61 DGRSEFAYGSFVRTVSLPVGADDDI 86

RA Shampula I.C., Baetian I., Fonteyne P.-A., Portaeis F.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the small heat shock protein (HSP20)
family.
DR EMBL: AF253462; AAM69224.1; -; Genomic_DNA.
DR InterPro: IPR002068; Hsp20.
DR Pfam: PF00011; HSP20; 1.
DR PROSITE: PS01031; HSP20; 1.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 86 AA; 9814 MW; 1D6966C08DE5845B CRC64;

Query Match 60.1%; Score 442; DB 2; Length 86;
Best Local Similarity 100.0%; Pred. No. 1.9e-32;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28 FAGLRPFPTDRLMLLEDEMEKGRYEVRAELPGVDPDKVDIMVRDGLTIKARTEQKDF 87
DB 1 FAGLRPFPTDRLMLLEDEMEKGRYEVRAELPGVDPDKVDIMVRDGLTIKARTEQKDF 60
QY 88 DGRSEFAYGSFVRTVSLPVGADDDI 113
DB 61 DGRSEFAYGSFVRTVSLPVGADDDI 86

RESULT 8
Q8KY76_MYCGN
ID Q8KY76_MYCGN PRELIMINARY; PRT; 86 AA.

AC Q8KY76;
DT 01-OCT-2002 (Tremblrel. 22, Created)
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Alpha-crystallin protein (Fragment).

GN Name=acr;
OS Mycobacterium genavense.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.

NCBI_TaxID=36812;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=M95-1325R;
RA Shampula I.C., Baetian I., Fonteyne P.-A., Portaeis F.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: Belongs to the small heat shock protein (HSP20)
family.
DR EMBL: AF253462; AAM69218.1; -; Genomic_DNA.
DR InterPro: IPR002068; Hsp20.
DR Pfam: PF00011; HSP20; 1.
DR PROSITE: PS01031; HSP20; 1.

FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 86 AA; 9814 MW; 1D6966C08DE5845B CRC64;

Query Match 60.1%; Score 442; DB 2; Length 86;
Best Local Similarity 100.0%; Pred. No. 1.9e-32;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28 FAGLRPFPTDRLMLLEDEMEKGRYEVRAELPGVDPDKVDIMVRDGLTIKARTEQKDF 87
DB 1 FAGLRPFPTDRLMLLEDEMEKGRYEVRAELPGVDPDKVDIMVRDGLTIKARTEQKDF 60
QY 88 DGRSEFAYGSFVRTVSLPVGADDDI 113
DB 61 DGRSEFAYGSFVRTVSLPVGADDDI 86

RESULT 9
Q8KY72_MYCCE
ID Q8KY72_MYCCE PRELIMINARY; PRT; 86 AA.

AC Q8KY72;
DT 01-OCT-2002 (Tremblrel. 22, Created)
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)

DE Alpha-crystallin protein (Fragment).

GN Name=acr;
OS Mycobacterium celatum.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.

NCBI_TaxID=28045;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=M95-143;
RA Shampula I.C., Baetian I., Fonteyne P.-A., Portaeis F.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: Belongs to the small heat shock protein (HSP20)
family.
DR EMBL: AF253460; AAM69222.1; -; Genomic_DNA.
DR InterPro: IPR002068; Hsp20.
DR Pfam: PF00011; HSP20; 1.
DR PROSITE: PS01031; HSP20; 1.

FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 86 AA; 9814 MW; 1D6966C08DE5845B CRC64;

Query Match 60.1%; Score 442; DB 2; Length 86;
Best Local Similarity 100.0%; Pred. No. 1.9e-32;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28 FAGLRPFPTDRLMLLEDEMEKGRYEVRAELPGVDPDKVDIMVRDGLTIKARTEQKDF 87
DB 1 FAGLRPFPTDRLMLLEDEMEKGRYEVRAELPGVDPDKVDIMVRDGLTIKARTEQKDF 60
QY 88 DGRSEFAYGSFVRTVSLPVGADDDI 113
DB 61 DGRSEFAYGSFVRTVSLPVGADDDI 86

RESULT 10
Q8KY68_MYCCH
ID Q8KY68_MYCCH PRELIMINARY; PRT; 86 AA.

AC Q8KY68;
DT 01-OCT-2002 (Tremblrel. 22, Created)
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Alpha-crystallin protein (Fragment).

GN Name=acr;
OS Mycobacterium chelonae.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.

NCBI_TaxID=1774;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=M8223;
RA Shampula I.C., Baetian I., Fonteyne P.-A., Portaeis F.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: Belongs to the small heat shock protein (HSP20)
family.
DR EMBL: AF253464; AAM69226.1; -; Genomic_DNA.
DR InterPro: IPR002068; Hsp20.
DR Pfam: PF00011; HSP20; 1.
DR PROSITE: PS01031; HSP20; 1.

FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 86 AA; 9800 MW; 0FF966C08DE5845B CRC64;

Query Match 59.9%; Score 441; DB 2; Length 86;
Best Local Similarity 98.8%; Pred. No. 2.4e-32;
Matches 85; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 28 FAGLRPFPTDRLMLLEDEMEKGRYEVRAELPGVDPDKVDIMVRDGLTIKARTEQKDF 87
DB 1 FAGLRPFPTDRLMLLEDEMEKGRYEVRAELPGVDPDKVDIMVRDGLTIKARTEQKDF 60
QY 88 DGRSEFAYGSFVRTVSLPVGADDDI 113
DB 61 DGRSEFAYGSFVRTVSLPVGADDDI 86

```
RESULT 11
08KY75 MYCPA PRELIMINARY; PRT; 86 AA.
ID 08KY75 MYCPA PRELIMINARY; PRT; 86 AA.
AC 08KY75 MYCPA PRELIMINARY; PRT; 86 AA.
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Alpha-crystallin protein (Fragment).
GN Name:acr;
OS Mycobacterium paratuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium avium complex (MAC).
OX NCBI_TaxID=1770;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=M2672;
RA Shamputa I.C., Bastian I., Fonteyne P.-A., Portaeis F.;
RL Submitted (APR-2000) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: Belongs to the small heat shock protein (HSP20)
family.
DR EMBL; AF253457; AAM69219.1; -; Genomic_DNA.
DR InterPro; IPR002068; Hsp20.
DR Pfam; PF00011; HSP20; 1.
DR PROSITE; PS01031; HSP20; 1.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 86 AA; 9800 MW; 0FF966C08DB5845B CRC64;

Query Match 59.9%; Score 441; DB 2; Length 86;
Best Local Similarity 98.8%; Pred. No. 2.4e-32;
Matches 85; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 28 FAGLRPTFDTRLMLBEDMKRGYEVRAELPGVDPDKVDIMVRDGLTIKAERTEQKF 87
Db 1 FAGLRPTFDTRLMLBEDMKRGYEVRAELPGVDPDKVDIMVRDGLTIKAERTEQKF 60

Qy 88 DGRSEFAYGSFVRTVSLPVGADDDI 113
Db 61 DGRSEFAYGSFVRTVSLPVGADDDV 86

RESULT 12
08KY69 MYCFO PRELIMINARY; PRT; 86 AA.
ID 08KY69 MYCFO PRELIMINARY; PRT; 86 AA.
AC 08KY69 MYCFO PRELIMINARY; PRT; 86 AA.
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Alpha-crystallin protein (Fragment).
GN Name:acr;
OS Mycobacterium fortuitum.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1766;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=M2672;
RA Shamputa I.C., Bastian I., Fonteyne P.-A., Portaeis F.;
RL Submitted (APR-2000) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: Belongs to the small heat shock protein (HSP20)
family.
DR EMBL; AF253453; AAM69225.1; -; Genomic_DNA.
DR InterPro; IPR002068; Hsp20.
DR Pfam; PF00011; HSP20; 1.
DR PROSITE; PS01031; HSP20; 1.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 86 AA; 9770 MW; B76666C0942C5442 CRC64;

Query Match 59.0%; Score 434; DB 2; Length 86;
Best Local Similarity 97.7%; Pred. No. 1e-31;
```

```
Matches 84; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 28 FAGLRPTFDTRLMLBEDMKRGYEVRAELPGVDPDKVDIMVRDGLTIKAERTEQKF 87
Db 1 FAGLRPTFDTRLMLBEDMKRGYEVRAELPGVDPDKVDIMVRDGLTIKAERTEQKF 60

Qy 88 DGRSEFAYGSFVRTVSLPVGADDDI 113
Db 61 DGRSEFAYGSFVRTVSLPVGADDDV 86

RESULT 13
08KY67 MYCXB PRELIMINARY; PRT; 86 AA.
ID 08KY67 MYCXB PRELIMINARY; PRT; 86 AA.
AC 08KY67 MYCXB PRELIMINARY; PRT; 86 AA.
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Alpha-crystallin-small heat shock protein (Fragment).
OS Mycobacterium xenopi.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1789;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=M2672;
RA Shamputa I.C., Bastian I., Fonteyne P.-A., Portaeis F.;
RL Submitted (APR-2000) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: Belongs to the small heat shock protein (HSP20)
family.
DR EMBL; AF261078; AAM69227.1; -; Genomic_DNA.
DR GO; GO:0006457; P:protein folding; IEA.
DR GO; GO:0006986; P:response to unfolded protein; IEA.
DR InterPro; IPR002068; Hsp20.
DR Pfam; PF00011; HSP20; 1.
DR PROSITE; PS01031; HSP20; 1.
KM Heat shock.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 86 AA; 9757 MW; BE9BD0D0BAC84E3 CRC64;

Query Match 49.3%; Score 363; DB 2; Length 86;
Best Local Similarity 79.1%; Pred. No. 2.8e-25;
Matches 68; Conservative 8; Mismatches 10; Indels 0; Gaps 0;

Qy 28 FAGLRPTFDTRLMLBEDMKRGYEVRAELPGVDPDKVDIMVRDGLTIKAERTEQKF 87
Db 1 FTGLHPIFDRLMLBEDMKRGYEVRAELPGVDPDKVDIMVRDGLTIKAERTEQKF 60

Qy 88 DGRSEFAYGSFVRTVSLPVGADDDI 113
Db 61 DGRSEFAYGSFVRTVSLPVGADDDV 86

RESULT 14
08KY84 MYCNO PRELIMINARY; PRT; 86 AA.
ID 08KY84 MYCNO PRELIMINARY; PRT; 86 AA.
AC 08KY84 MYCNO PRELIMINARY; PRT; 86 AA.
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Alpha-crystallin protein (Fragment).
GN Name:acr;
OS Mycobacterium nonchromogenicum.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1782;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC19530;
RA Shamputa I.C., Bastian I., Fonteyne P.-A., Portaeis F.;
RL Submitted (APR-2000) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: Belongs to the small heat shock protein (HSP20)
family.
```

DR EMBL; AF253445; AAM69207.1; -; Genomic_DNA.
 DR InterPro; IPR02068; HSP20.
 DR Pfam; PF00011; HSP20; 1.
 DR PROSITE; PS01031; HSP20; 1.
 FT NON_TER 1
 SQ SEQUENCE 86 AA; 9722 MW; BAF085B56BF1DBF3 CRC64;

Query Match 46.6%; Score 343; DB 2; Length 86;
 Best Local Similarity 74.4%; Pred. No. 1.9e-23;
 Matches 64; Conservative 11; Mismatches 11; Indels 0; Gaps 0;

QY 28 FAGLRPTFTRLMLLEDEMKGRYEVRALPGVDPDKVDIMVRDGLTIKAERTQKDF 87
 DB 1 FAGLRPAFDSRLMLLEDEMKGRYELRAIPGVDPKDIETVLDGQLTIKAERSEKKEF 60

QY 88 DGRSEFAYGSFVRTVSLPGVADDDI 113
 DB 61 NGRSEFCYGSFRTVPLPAGVTEDGI 86

RESULT 15

Q8KY71 MYCGS PRELIMINARY; PRT; 86 AA.
 AC Q8KY71;
 DT 01-OCT-2002 (TREMBlrel. 22, Created)
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Alpha-crystallin protein (Fragment).
 GN Name=acr;
 OS Mycobacterium gastri.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_Taxid=1777;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=M97-938;
 RA Shampita I.C., Bastian I., Fonteyne P.-A., Portaeis F.;
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: Belongs to the small heat shock protein (HSP20) family.
 DR EMBL; AF253461; AAM69223.1; -; Genomic_DNA.
 DR InterPro; IPR02068; HSP20.
 DR Pfam; PF00011; HSP20; 1.
 DR PROSITE; PS01031; HSP20; 1.
 FT NON_TER 1
 SQ SEQUENCE 86 AA; 9722 MW; BAF085B56BF1DBF3 CRC64;

Query Match 46.6%; Score 343; DB 2; Length 86;
 Best Local Similarity 74.4%; Pred. No. 1.9e-23;
 Matches 64; Conservative 11; Mismatches 11; Indels 0; Gaps 0;

QY 28 FAGLRPTFTRLMLLEDEMKGRYEVRALPGVDPDKVDIMVRDGLTIKAERTQKDF 87
 DB 1 FAGLRPAFDSRLMLLEDEMKGRYELRAIPGVDPKDIETVLDGQLTIKAERSEKKEF 60

QY 88 DGRSEFAYGSFVRTVSLPGVADDDI 113
 DB 61 NGRSEFCYGSFRTVPLPAGVTEDGI 86

Search completed: March 23, 2006, 05:24:22
 Job time : 38.8298 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using bw model

Run on: March 23, 2006, 05:24:53 ; Search time 9.82128 Seconds

(without alignments)
1212.194 Million cell updates/sec

Title: US-10-617-038-24

Perfect score: 736
Sequence: 1 MATLPPVQRHRSIFPERSE.....SVAVSEKPTKHIQIRSTN 144

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents, AA:*

- 1: /cgn2_6/ptodata/1/1aa/5_COMB.pep:*
- 2: /cgn2_6/ptodata/1/1aa/6_COMB.pep:*
- 3: /cgn2_6/ptodata/1/1aa/H_COMB.pep:*
- 4: /cgn2_6/ptodata/1/1aa/PCUTS_COMB.pep:*
- 5: /cgn2_6/ptodata/1/1aa/RE_COMB.pep:*
- 6: /cgn2_6/ptodata/1/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	736	100.0	144	2	US-09-462-556-2
2	736	100.0	331	2	US-09-223-040-4
3	736	100.0	331	2	US-09-287-849-4
4	736	100.0	368	2	US-09-287-849-24
5	736	100.0	433	2	US-09-287-849-14
6	736	100.0	856	2	US-09-287-849-12
7	718	97.6	143	2	US-09-461-774-8
8	745	93.3	48	2	US-09-157-689-18
9	245	33.3	48	2	US-08-447-398-18
10	245	33.3	48	2	US-09-953-510-18
11	146.5	17.9	151	2	US-09-902-540-13672
12	132	13.9	139	2	US-09-107-532A-4324
13	116	15.8	131	2	US-09-902-540-11500
14	100.5	13.7	250	2	US-09-949-016-8090
15	94	12.8	156	2	US-09-543-681A-7799
16	93.5	12.7	167	2	US-09-835-909A-2
17	84.5	11.5	759	2	US-09-792-024-76
18	83.5	11.3	177	2	US-10-105-427-2
19	82.5	11.2	274	2	US-09-107-532A-4705
20	81	11.0	460	2	US-09-199-637A-132
21	79	10.7	175	2	US-09-513-999C-7967
22	79	10.7	183	2	US-09-949-016-8416
23	79	10.7	250	2	US-09-513-999C-7968
24	79	10.7	918	2	US-09-198-452A-1072
25	79	10.7	918	2	US-09-438-185A-999
26	78.5	10.7	475	2	US-09-252-991A-29494
27	75.5	10.3	199	1	US-08-900-407-4

28	75.5	10.3	240	2	US-09-949-016-10863	Sequence 10863, A
29	75.5	10.3	459	2	US-09-513-783A-170	Sequence 170, App
30	75.5	10.3	459	2	US-10-100-957A-170	Sequence 170, App
31	74.5	10.1	168	2	US-09-949-016-8494	Sequence 8494, Ap
32	72.5	9.9	291	2	US-09-902-540-12550	Sequence 12550, A
33	72.5	9.9	545	2	US-09-902-540-10533	Sequence 10533, A
34	72	9.8	188	2	US-09-902-540-9929	Sequence 9929, Ap
35	72	9.8	332	2	US-09-605-703B-2268	Sequence 2268, Ap
36	71.5	9.7	193	2	US-08-900-407-3	Sequence 3, Appli
37	71.5	9.7	311	2	US-09-543-681A-5563	Sequence 5563, Ap
38	71.5	9.7	693	2	US-09-376-343-2	Sequence 2, Appli
39	71	9.6	497	2	US-09-328-352-7286	Sequence 7286, Ap
40	70.5	9.6	923	2	US-09-328-352-4371	Sequence 4371, Ap
41	70.5	9.6	1132	2	US-09-248-796A-15026	Sequence 15026, A
42	68.5	9.3	167	2	US-09-489-039A-13896	Sequence 13896, A
43	68.5	9.3	486	2	US-09-252-991A-24336	Sequence 24336, A
44	68.5	9.3	594	2	US-09-252-991A-18099	Sequence 18099, A
45	68	9.2	248	2	US-09-489-039A-11767	Sequence 11767, A

ALIGNMENTS

```

RESULT 1
US-09-462-556-2
; Sequence 2, Application US/09462556
; Patent No. 6403100
; GENERAL INFORMATION:
; APPLICANT: Barry III, Clifton E.
; APPLICANT: Yuan, Ying
; APPLICANT: Crane, Deborah D.
; APPLICANT: The Government of the United States of America
; APPLICANT: as represented by The Secretary of the
; APPLICANT: Department of Health and Human Services
; TITLE OF INVENTION: Method of Attenuating Pathogenic Mycobacteria and
; TITLE OF INVENTION: Strains of Mycobacteria So Attenuated
; FILE REFERENCE: 015280-333100US
; CURRENT APPLICATION NUMBER: US/09/462,556
; PRIOR FILING DATE: 2000-06-27
; PRIOR APPLICATION NUMBER: US 60/052,199
; PRIOR FILING DATE: 1997-07-10
; PRIOR APPLICATION NUMBER: WO PCT/US98/14227
; PRIOR FILING DATE: 1998-07-09
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 144
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: alpha-crystallin heat shock protein
US-09-462-556-2
Query Match 100.0%; Score 736; DB 2; Length 144;
Best Local Similarity 100.0%; Pred. No. 9.8e-81;
Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MATLPPVQRHRSIFPERSEIFAAPSPAGIRPFDFRIMLEBEMKRGREVAELPGV 60
DB 1 MATLPPVQRHRSIFPERSEIFAAPSPAGIRPFDFRIMLEBEMKRGREVAELPGV 60
QY 61 DPDDVDIMVWDGQLTIAKRTKQDFGRSEFVYGSFVTVSLPVGADDDIDKATYDKG 120
DB 61 DPDDVDIMVWDGQLTIAKRTKQDFGRSEFVYGSFVTVSLPVGADDDIDKATYDKG 120
QY 121 ILTVSVAVSEKPTKHIQIRSTN 144
DB 121 ILTVSVAVSEKPTKHIQIRSTN 144
RESULT 2
US-09-223-040-4
; Sequence 4, Application US/09223040

```

```
Patent No. 6544522
GENERAL INFORMATION:
APPLICANT: Skeiky, Yasir
APPLICANT: Alderson, Mark
APPLICANT: Campos-Neto, Antonio
APPLICANT: Corixa Corporation
TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens
FILE REFERENCE: 014058-009010US
CURRENT FILING DATE: 1998-12-30
NUMBER OF SEQ ID NOS: 10
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 4
LENGTH: 331
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: tri-fusion
US-09-223-040-4
```

```
Query Match          100.0%; Score 736; DB 2; Length 331;
Best Local Similarity 100.0%; Pred. No. 3,4e-80;
Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MATLLPVQRHPRSLFPEFSELFAFPSPAGLRPTFDTRLMLRLEDEMKEGRYEVRALPGV 60
   |||||
DB 8 MATLLPVQRHPRSLFPEFSELFAFPSPAGLRPTFDTRLMLRLEDEMKEGRYEVRALPGV 67
   |||||
QY 61 DDDKDVIMVRDGLTIKARTEQKDFGRSEFAYGSFVRTVSLPVGADEDDIKATYDKG 120
   |||||
DB 68 DDDKDVIMVRDGLTIKARTEQKDFGRSEFAYGSFVRTVSLPVGADEDDIKATYDKG 127
   |||||
QY 121 ILTVSAVSESGKPTKEKHQIRSTN 144
   |||||
DB 128 ILTVSAVSESGKPTKEKHQIRSTN 151
   |||||
```

```
RESULT 3
US-09-287-849-4
Sequence 4, Application US/09287849
Patent No. 6627198
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Dillon, Davin C.
APPLICANT: Alderson, Mark
APPLICANT: Campos-Neto, Antonio
APPLICANT: Corixa Corporation
TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens
FILE REFERENCE: 014058-009020US
CURRENT FILING DATE: 1999-04-07
CURRENT FILING DATE: 1999-04-07
PRIOR APPLICATION NUMBER: US 08/818,112
PRIOR FILING DATE: 1997-03-13
PRIOR APPLICATION NUMBER: US 08/942,578
PRIOR FILING DATE: 1997-10-01
PRIOR APPLICATION NUMBER: US 09/025,197
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 09/056,556
PRIOR FILING DATE: 1998-04-07
PRIOR APPLICATION NUMBER: US 09/223,040
PRIOR FILING DATE: 1998-12-30
NUMBER OF SEQ ID NOS: 46
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 4
LENGTH: 331
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: tri-fusion
US-09-287-849-4
```

```
Query Match          100.0%; Score 736; DB 2; Length 331;
Best Local Similarity 100.0%; Pred. No. 3,4e-80;
Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MATLLPVQRHPRSLFPEFSELFAFPSPAGLRPTFDTRLMLRLEDEMKEGRYEVRALPGV 60
   |||||
DB 8 MATLLPVQRHPRSLFPEFSELFAFPSPAGLRPTFDTRLMLRLEDEMKEGRYEVRALPGV 67
   |||||
QY 61 DDDKDVIMVRDGLTIKARTEQKDFGRSEFAYGSFVRTVSLPVGADEDDIKATYDKG 120
   |||||
DB 68 DDDKDVIMVRDGLTIKARTEQKDFGRSEFAYGSFVRTVSLPVGADEDDIKATYDKG 127
   |||||
QY 121 ILTVSAVSESGKPTKEKHQIRSTN 144
   |||||
DB 128 ILTVSAVSESGKPTKEKHQIRSTN 151
   |||||
```

```
RESULT 4
US-09-287-849-24
Sequence 24, Application US/09287849
Patent No. 6627198
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Dillon, Davin C.
APPLICANT: Alderson, Mark
APPLICANT: Campos-Neto, Antonio
APPLICANT: Corixa Corporation
TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens
FILE REFERENCE: 014058-009020US
CURRENT FILING DATE: 1999-04-07
CURRENT FILING DATE: 1999-04-07
PRIOR APPLICATION NUMBER: US 08/818,112
PRIOR FILING DATE: 1997-03-13
PRIOR APPLICATION NUMBER: US 08/942,578
PRIOR FILING DATE: 1997-10-01
PRIOR APPLICATION NUMBER: US 09/025,197
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 09/056,556
PRIOR FILING DATE: 1998-04-07
PRIOR APPLICATION NUMBER: US 09/223,040
PRIOR FILING DATE: 1998-12-30
NUMBER OF SEQ ID NOS: 46
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 24
LENGTH: 368
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: tri-fusion
US-09-287-849-24
```

```
Query Match          100.0%; Score 736; DB 2; Length 368;
Best Local Similarity 100.0%; Pred. No. 4e-80;
Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MATLLPVQRHPRSLFPEFSELFAFPSPAGLRPTFDTRLMLRLEDEMKEGRYEVRALPGV 60
   |||||
DB 9 MATLLPVQRHPRSLFPEFSELFAFPSPAGLRPTFDTRLMLRLEDEMKEGRYEVRALPGV 68
   |||||
QY 61 DDDKDVIMVRDGLTIKARTEQKDFGRSEFAYGSFVRTVSLPVGADEDDIKATYDKG 120
   |||||
DB 69 DDDKDVIMVRDGLTIKARTEQKDFGRSEFAYGSFVRTVSLPVGADEDDIKATYDKG 128
   |||||
QY 121 ILTVSAVSESGKPTKEKHQIRSTN 144
   |||||
DB 129 ILTVSAVSESGKPTKEKHQIRSTN 152
   |||||
```

```
RESULT 5
US-09-287-849-14
```

```
; Sequence 14, Application US/09287849
; Patent No. 6627198
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Alderson, Mark
; APPLICANT: Campos-Neco, Antonio
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens
; TITLE OF INVENTION: and Their Uses
; FILE REFERENCE: 014058-009020US
; CURRENT APPLICATION NUMBER: US/09/287,849
; PRIOR FILING DATE: 1999-04-07
; PRIOR APPLICATION NUMBER: US 08/818,112
; PRIOR FILING DATE: 1997-03-13
; PRIOR APPLICATION NUMBER: US 08/942,578
; PRIOR FILING DATE: 1997-10-01
; PRIOR APPLICATION NUMBER: US 09/025,197
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 09/056,556
; PRIOR FILING DATE: 1998-04-07
; PRIOR APPLICATION NUMBER: US 09/223,040
; PRIOR FILING DATE: 1998-12-30
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 433
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:tetra-fusion
US-09-287-849-14
```

```
Query Match 100.0%; Score 736; DB 2; Length 433;
Best Local Similarity 100.0%; Pred. No. 5.2e-80;
Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MATTLPVQRHRSLEPPESELPFAFPSPAGLRPTFTDRLMLRLEDEMKRGYEVRAELPGV 60
Db 9 MATTLPVQRHRSLEPPESELPFAFPSPAGLRPTFTDRLMLRLEDEMKRGYEVRAELPGV 68

Qy 61 DPDKVDIMVRDGLTTIKARTEQKDPGRSEFAVGSFVRTVSLPVGADEDDIKATYDKG 120
Db 69 DPDKVDIMVRDGLTTIKARTEQKDPGRSEFAVGSFVRTVSLPVGADEDDIKATYDKG 128

Qy 121 LITVSAVSEGGKPTKEKHIIQIRSTN 144
Db 129 LITVSAVSEGGKPTKEKHIIQIRSTN 152
```

```
RESULT 6
US-09-287-849-12
; Sequence 12, Application US/09287849
; Patent No. 6627198
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Alderson, Mark
; APPLICANT: Campos-Neco, Antonio
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens
; TITLE OF INVENTION: and Their Uses
; FILE REFERENCE: 014058-009020US
; CURRENT APPLICATION NUMBER: US/09/287,849
; PRIOR FILING DATE: 1999-04-07
; PRIOR APPLICATION NUMBER: US 08/818,112
; PRIOR FILING DATE: 1997-03-13
; PRIOR APPLICATION NUMBER: US 08/942,578
; PRIOR FILING DATE: 1997-10-01
; PRIOR APPLICATION NUMBER: US 09/025,197
; PRIOR FILING DATE: 1998-02-18
```

```
; PRIOR APPLICATION NUMBER: US 09/056,556
; PRIOR FILING DATE: 1998-04-07
; PRIOR APPLICATION NUMBER: US 09/223,040
; PRIOR FILING DATE: 1998-12-30
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 856
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:penta-fusion
US-09-287-849-12
```

```
Query Match 100.0%; Score 736; DB 2; Length 856;
Best Local Similarity 100.0%; Pred. No. 1.4e-79;
Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MATTLPVQRHRSLEPPESELPFAFPSPAGLRPTFTDRLMLRLEDEMKRGYEVRAELPGV 60
Db 9 MATTLPVQRHRSLEPPESELPFAFPSPAGLRPTFTDRLMLRLEDEMKRGYEVRAELPGV 68

Qy 61 DPDKVDIMVRDGLTTIKARTEQKDPGRSEFAVGSFVRTVSLPVGADEDDIKATYDKG 120
Db 69 DPDKVDIMVRDGLTTIKARTEQKDPGRSEFAVGSFVRTVSLPVGADEDDIKATYDKG 128

Qy 121 LITVSAVSEGGKPTKEKHIIQIRSTN 144
Db 129 LITVSAVSEGGKPTKEKHIIQIRSTN 152
```

```
RESULT 7
US-09-461-774-8
; Sequence 8, Application US/09461774
; Patent No. 6887481
; GENERAL INFORMATION:
; APPLICANT: CHAN, Lily
; APPLICANT: CHUNG, Maxey Ching Ming
; APPLICANT: LIM, Renee Lay Hong
; TITLE OF INVENTION: Bacterial-derived molecules and therapeutic and
; TITLE OF INVENTION: diagnostic uses therefor
; FILE REFERENCE: 1781-0180P
; CURRENT APPLICATION NUMBER: US/09/461,774
; PRIOR FILING DATE: 1999-12-15
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 143
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-09-461-774-8
```

```
Query Match 97.6%; Score 718; DB 2; Length 143;
Best Local Similarity 98.6%; Pred. No. 1.4e-78;
Matches 141; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy 2 ATTLPVQRHRSLEPPESELPFAFPSPAGLRPTFTDRLMLRLEDEMKRGYEVRAELPGV 61
Db 1 ATTLPVQRHRSLEPPESELPFAFPSPAGLRPTFTDRLMLRLEDEMKRGYEVRAELPGV 60

Qy 62 PDKVDIMVRDGLTTIKARTEQKDPGRSEFAVGSFVRTVSLPVGADEDDIKATYDKGI 121
Db 61 PDKVDIMVRDGLTTIKARTEQKDPGRSEFAVGSFVRTVSLPVGADEDDIKATYDKGI 120

Qy 122 LITVSAVSEGGKPTKEKHIIQIRSTN 144
Db 121 LITVSAVSEGGKPTKEKHIIQIRSTN 143
```

```
RESULT 8
US-09-157-689-18
; Sequence 18, Application US/09157689
; Patent No. 6599510
```

GENERAL INFORMATION:
APPLICANT: Horwitz, Marcus A.
TITLE OF INVENTION: Abundant Extracellular Products and Methods for Their Production and Use
NUMBER OF SEQUENCES: 95
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kurt A. Maclean
STREET: 2029 Century Park East, Suite 3800
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90067
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/157,689
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/568,357
FILING DATE: 06-DEC-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/551,149
FILING DATE: 31-OCT-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/447,398
FILING DATE: 23-MAY-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/289,667
FILING DATE: 12-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/156,358
FILING DATE: 23-NOV-1993
ATTORNEY/AGENT INFORMATION:
NAME: Maclean, Kurt A.
REGISTRATION NUMBER: 31,118
REFERENCE/DOCKET NUMBER: 118-119
TELECOMMUNICATION INFORMATION:
TELEPHONE: (310) 788-5000
TELEFAX: (310) 277-1297
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 48 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Mycobacterium tuberculosis
STRAIN: Erdman
US-09-157-689-18

Query Match 33.3%; Score 245; DB 2; Length 48;
Best Local Similarity 100.0%; Pred. No. 2.9e-22;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 35 FDTRLMLBDEMKRGYEVRAELPGVDPDKVDIMVRDGLTIKAERT 82
Db 1 FDTRLMLBDEMKRGYEVRAELPGVDPDKVDIMVRDGLTIKAERT 48

RESULT 9
US-08-447-398-18
Sequence 18, Application US/08447398
Patent No. 6761894
GENERAL INFORMATION:

APPLICANT: Horwitz, Marcus A.
TITLE OF INVENTION: Abundant Extracellular Products and Methods for Their Production and Use
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kurt A. Maclean
STREET: 2029 Century Park East, Suite 3800
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90067
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/447,398
FILING DATE: 23-MAY-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/289,667
FILING DATE: 12-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/156,358
FILING DATE: 23-NOV-1993
ATTORNEY/AGENT INFORMATION:
NAME: Maclean, Kurt A.
REGISTRATION NUMBER: 31,118
REFERENCE/DOCKET NUMBER: 111-161
TELECOMMUNICATION INFORMATION:
TELEPHONE: (310) 788-5000
TELEFAX: (310) 277-1297
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 48 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Mycobacterium tuberculosis
STRAIN: Erdman
US-08-447-398-18

Query Match 33.3%; Score 245; DB 2; Length 48;
Best Local Similarity 100.0%; Pred. No. 2.9e-22;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 35 FDTRLMLBDEMKRGYEVRAELPGVDPDKVDIMVRDGLTIKAERT 82
Db 1 FDTRLMLBDEMKRGYEVRAELPGVDPDKVDIMVRDGLTIKAERT 48

RESULT 10
US-09-953-510-18
Sequence 18, Application US/09953510
Patent No. 6818223
GENERAL INFORMATION:
APPLICANT: Horwitz, Marcus A.
TITLE OF INVENTION: Abundant Extracellular Products and Methods for Their Production and Use
NUMBER OF SEQUENCES: 91
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kurt A. Maclean
STREET: 2029 Century Park East, Suite 3800
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90067

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0,
Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/953,510
FILING DATE: 14-Sep-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/447,398
FILING DATE: 23-MAY-1995
APPLICATION NUMBER: US 08/289,667
FILING DATE: 12-AUG-1994
APPLICATION NUMBER: US 08/156,358
FILING DATE: 23-NOV-1993
ATTORNEY/AGENT INFORMATION:
NAME: MacLean, Kurt A.
REGISTRATION NUMBER: 31,118
REFERENCE/DOCKET NUMBER: 112-272
TELECOMMUNICATION INFORMATION:
TELEPHONE: (310) 788-5000
TELEFAX: (310) 277-1297
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 48 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Mycobacterium tuberculosis
STRAIN: Erdman
SEQUENCE DESCRIPTION: SEQ ID NO: 18:
US-09-953-510-18
Query Match 33.3%; Score 245; DB 2; Length 48;
Best Local Similarity 100.0%; Pred. No. 2.9e-22;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 35 FDTFLMLEDMEKGRYEVRAELPGVDPDKVDIMVRDGLTIKAERT 82
Db 1 FDTFLMLEDMEKGRYEVRAELPGVDPDKVDIMVRDGLTIKAERT 48
RESULT 11
US-09-902-540-13672
Sequence 13672, Application US/09902540
Patent No. 6833447
GENERAL INFORMATION:
APPLICANT: Goldman, Barry S.
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Wiegand, Roger C.
TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
FILE REFERENCE: 38-10(115849)B
CURRENT FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: 60/4217,883
PRIOR FILING DATE: 2000-07-10
NUMBER OF SEQ ID NOS: 16825
SEQ ID NO 13672
LENGTH: 151
TYPE: PRT
ORGANISM: Myxococcus xanthus
US-09-902-540-13672
Query Match 19.9%; Score 146.5; DB 2; Length 151;
Best Local Similarity 33.1%; Pred. No. 1.2e-09;

Matches 39; Conservative 29; Mismatches 31; Indels 19; Gaps 6;
Qy 30 GLRPTFDTFLMLEDMEKGR--YEVRAELPGVDPDKVDIMVRDGLTIKAERTQKDF 87
Db 35 GLRPTDF-----EVKRTQDAFIRKADVPGE-EDDLITLAEKRLTISGRREERD 84
Qy 88 DGRSEFA---YGSFVATYSLPVGADEDDIKATYDKILTVSV-ANSEKPTKHIOI 140
Db 85 EGRRYVAYERNYGSFSRTFTLPGRVANVDVQADFKSGVLNVRIPKSEQOP--KRIVY 140
RESULT 12
US-09-107-532A-4324
Sequence 4324, Application US/09107532A
Patent No. 6583275
GENERAL INFORMATION:
APPLICANT: Lynn A Doucetere-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Walham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Ariadello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781) 893-5007
TELEFAX: (781) 893-8277
INFORMATION FOR SEQ ID NO: 4324:
SEQUENCE CHARACTERISTICS:
LENGTH: 139 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...139
SEQUENCE DESCRIPTION: SEQ ID NO: 4324:
US-09-107-532A-4324
Query Match 17.9%; Score 132; DB 2; Length 139;
Best Local Similarity 33.1%; Pred. No. 6.1e-08;
Matches 46; Conservative 26; Mismatches 47; Indels 20; Gaps 10;
Qy 12 RSLPFRSESLFA-AFSPFAGLR--PTFDTFLMLEDMEKGRYEVRAELPGVDPDKVDI 68
Db 10 RDMFPDNDVFSRAFNDFLVSSYPKVD-----LVENKE--YKLTADMFGCDKE-DTIV 61
Qy 69 MVRDGLTIYA---ERTEQDFDQ---RSEFAYGSFVATYSLPVGADEDDIKATYDKIL 122
Db 62 EYSDNTLTISANSHSHTEDEK-DGNVYRKERHSHSVYSKRSFYL-P-NVDEKITGTGKXGVL 119

```
Qy      123 TVSVAVSECKPTE-KHIQI 140
        : : : | | | : :
Db      120 KLVLPKTAQPKETKKIEL 138
```

```

RESULT 13
US-09-902-540-11500
: Sequence 11500, Application US/09902540
: Patent No. 683447
:
: GENERAL INFORMATION:
: APPLICANT: Goldman, Barry S.
: APPLICANT: Hinkle, Gregory J.
: APPLICANT: Slater, Steven C.
: APPLICANT: Wiegand, Roger C.
: TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
: FILE REFERENCE: 38-101(15849)B
:
: CURRENT APPLICATION NUMBER: US/09/902,540
: CURRENT FILING DATE: 2001-07-10
: PRIOR APPLICATION NUMBER: 60/217,883
: PRIOR FILING DATE: 2000-07-10
:
: NUMBER OF SEQ ID NOS: 16825
:
: SEQ ID NO 11500
:
: LENGTH: 131
:
: TYPE: prt
:
: ORGANISM: Myxococcus xanthus
:
: US-09-902-540-11500

```

```

RESULT 14
US-09-949-016-8090
Sequence 8090, Application US/09949016
Patent No. 681235
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL0010107
CURRENT APPLICATION NUMBER: US/09/949, 016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 8090
LENGTH: 250
TYPE: prt
ORGANISM: Human
US-09-949-016-8090

```

```

OY      119 KG 120
DB      192 LG 193

RESULT 15
US-09-543-681A-7799
; Sequence 7799, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 7799
; LENGTH: 156
; TYPE: prt
; ORGANISM: Proteus mirabilis
US-09-543-681A-7799

```

Search completed: March 23, 2006, 05:30:14
Job time : 10.9213 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: March 23, 2006, 06:14:33 ; Search time 30.2823 Seconds
(without alignments)
1986.885 Million cell updates/sec

Title: US-10-617-038-24

Perfect score: 736
Sequence: 1 MATTLPVQRHPRSLPFPSR.....SVAVSEKPTKHIQIRSTN 144

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_AA_Main:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
- 3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
- 4: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
- 6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	736	100.0	144	US-10-617-038-24	Sequence 24, Appl
2	736	100.0	144	US-10-332-512A-23	Sequence 23, Appl
3	736	100.0	331	US-09-287-849-4	Sequence 4, Appl
4	736	100.0	331	US-10-359-460-4	Sequence 4, Appl
5	736	100.0	331	US-10-359-459-4	Sequence 4, Appl
6	736	100.0	368	US-09-287-849-24	Sequence 24, Appl
7	736	100.0	368	US-10-359-460-24	Sequence 24, Appl
8	736	100.0	433	US-09-287-849-14	Sequence 14, Appl
9	736	100.0	433	US-10-359-460-14	Sequence 14, Appl
10	736	100.0	856	US-09-287-849-12	Sequence 12, Appl
11	736	100.0	856	US-10-359-460-12	Sequence 12, Appl
12	736	100.0	875	US-10-368-983-13	Sequence 13, Appl
13	245	33.3	48	US-09-953-510-18	Sequence 18, Appl
14	245	33.3	48	US-09-953-413-18	Sequence 18, Appl
15	245	33.3	48	US-10-147-255-18	Sequence 18, Appl
16	245	33.3	48	US-10-147-255-18	Sequence 18, Appl
17	184.5	25.1	175	US-10-732-923-21800	Sequence 21800, A
18	183.5	24.9	147	US-10-732-923-21778	Sequence 21778, A
19	173	23.5	173	US-10-732-923-21770	Sequence 21770, A
20	166.5	22.6	231	US-10-732-923-21756	Sequence 21756, A
21	165	22.4	248	US-10-732-923-21792	Sequence 21792, A
22	164.5	22.4	159	US-10-481-265-17	Sequence 17, Appl
23	160.5	21.8	185	US-10-732-923-21698	Sequence 21698, A
24	158.5	21.5	202	US-10-732-923-21766	Sequence 21766, A
25	156	21.2	146	US-10-047-260-14	Sequence 14, Appl
26	155.5	21.1	143	US-10-732-923-22005	Sequence 22005, A
27	155.5	21.1	169	US-10-732-923-21700	Sequence 21700, A

28	154.5	21.0	134	5	US-10-732-923-21697	Sequence 21697, A
29	154.5	21.0	175	5	US-10-732-923-21769	Sequence 21769, A
30	152.5	20.7	132	5	US-10-732-923-21763	Sequence 21763, A
31	151	20.5	168	5	US-10-732-923-21749	Sequence 21749, A
32	150.5	20.4	145	5	US-10-732-923-21749	Sequence 21749, A
33	150.5	20.4	169	5	US-10-732-923-21816	Sequence 21816, A
34	150.5	20.4	188	4	US-10-657-740-8	Sequence 8, Appl
35	150.5	20.4	188	4	US-10-732-923-21817	Sequence 21817, A
36	149.5	20.3	140	5	US-10-732-923-21759	Sequence 21759, A
37	148	20.1	140	5	US-10-732-923-21703	Sequence 21703, A
38	148	20.1	146	5	US-10-732-923-21755	Sequence 21755, A
39	146.5	19.9	215	5	US-10-732-923-21774	Sequence 21774, A
40	145.5	19.8	151	5	US-10-732-923-21774	Sequence 21774, A
41	145.5	19.8	159	5	US-10-732-923-21709	Sequence 21709, A
42	144.5	19.6	150	5	US-10-732-923-21853	Sequence 21853, A
43	144.5	19.6	166	5	US-10-732-923-21771	Sequence 21771, A
44	144	19.6	125	5	US-10-732-923-21760	Sequence 21760, A
45	142.5	19.4	145	5	US-10-732-923-21854	Sequence 21854, A

ALIGNMENTS

```
RESULT 1
US-10-617-038-24
; Sequence 24, Application US/10617038
; Publication No. US20040057963A1
; GENERAL INFORMATION:
; APPLICANT: Andersen, Peter
; APPLICANT: Rosenkrands, Ida
; APPLICANT: Scrym, Anette
; TITLE OF INVENTION: Therapeutic TB vaccine
; FILE REFERENCE: SSI5AUSA
; CURRENT APPLICATION NUMBER: US/10/617,038
; PRIOR FILING DATE: 2003-07-11
; PRIOR APPLICATION NUMBER: DK PA 2002 01098
; PRIOR FILING DATE: 2002-07-13
; PRIOR APPLICATION NUMBER: US 60/401,725
; PRIOR FILING DATE: 2002-08-07
; NUMBER OF SEQ ID NOS: 187
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 24
; LENGTH: 144
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-617-038-24

Query Match      100.0%; Score 736; DB 4; Length 144;
Best Local Similarity 100.0%; Pred. No. 1.1e-72;
Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MATTLPVQRHPRSLPFPSRFAAGLRPTDRLRLRDEMEKGRYVRAELRGV 60
DB      1 MATTLPVQRHPRSLPFPSRFAAGLRPTDRLRLRDEMEKGRYVRAELRGV 60

QY      61 DPKDQVIMRDQLTIKARTQKDFDGRSEFAYGSFVATVSLPVGADBDIKATYDKG 120
DB      61 DPKDQVIMRDQLTIKARTQKDFDGRSEFAYGSFVATVSLPVGADBDIKATYDKG 120

QY      121 ILTVSAVSEKPTKHIQIRSTN 144
DB      121 ILTVSAVSEKPTKHIQIRSTN 144

RESULT 2
US-10-332-512A-23
; Sequence 23, Application US/10332512A
; Publication No. US20040180056A1
; GENERAL INFORMATION:
; APPLICANT: ORME, Ian M.
; APPLICANT: BELISSE, John T.
; TITLE OF INVENTION: MID-LIFE VACCINE AND METHODS FOR BOOSTING ANTI-MYCOBACTERIAL IMMOT
; FILE REFERENCE: 38861-186292
```

```

: CURRENT APPLICATION NUMBER: US-10/332,512A
:
: CURRENT FILING DATE: 2003-01-10
: PRIOR APPLICATION NUMBER: PCT/US01/21717
: PRIOR FILING DATE: 2001-07-10
: PRIOR APPLICATION NUMBER: US 60/217,646
: PRIOR FILING DATE: 2000-07-10
: NUMBER OF SEQ ID NOS: 31
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 23
: LENGTH: 144
: TYPE: prt
: ORGANISM: Mycobacterium tuberculosis
: US-10-332-512A-23

```

[illegible]

```

RESULT 3
US-09-287-849-4
; Sequence 4, Application US/09287849
; Patent No. US2002009459A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yaelir A.W.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Alderson, Mark
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens
; TITLE OF INVENTION: and Their Uses
; FILE REFERENCE: 014058-009020US
; CURRENT APPLICATION NUMBER: US/09/287,849
; CURRENT FILING DATE: 1999-04-07
; PRIOR APPLICATION NUMBER: US 08/818,112
; PRIOR FILING DATE: 1997-03-13
; PRIOR APPLICATION NUMBER: US 08/942,578
; PRIOR FILING DATE: 1997-10-01
; PRIOR APPLICATION NUMBER: US 09/025,197
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 09/056,556
; PRIOR FILING DATE: 1998-04-07
; PRIOR APPLICATION NUMBER: US 09/223,040
; PRIOR FILING DATE: 1998-12-30
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 331
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:tri-fusion
; US-09-287-849-4

```

Query Match	100.0%	Score 736;	DB 3;	Length 331;
Best Local Similarity	100.0%	Pred. No. 3.5e-72;		
Matches 144;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MATTLPGQRNRSLPPEPSELPAAPPSAGCRPTFTDTLMLLEDEMGKREYVAELPGV	60	

Db	8	MAITLPGVQHPRSLPEFSESLPAAPSPFAGLRPTDTRLKMLBDMKREGRJEVAAELPGV	67
QY	61	DPKDDVDIMRWQDQLTIKARTEQNDPGGRSEFAAGSVPTVSLPVGADDEDDIKATYTDKG	120
Db	68	DPKDDVDIMRWQDQLTIKARTEQNDPGGRSEFAAGSVPTVSLPVGADDEDDIKATYTDKG	127
QY	121	ILTVSAVSEGGKPTENHIQIRSTN	144
Db	128	ILTVSAVSEGGKPTENHIQIRSTN	151

```

RESULT 4
US-10-359-460-4
; Sequence 4, Application US/10359460
; Publication No. US20030147911A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skelky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Alderson, Mark
; APPLICANT: Campos-Neco, Antonio
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens
; TITLE OF INVENTION: and Their Uses
; FILE REFERENCE: 014058-009020US
; CURRENT APPLICATION NUMBER: US/10/359,460
; CURRENT FILING DATE: 2003-02-05
; PRIOR APPLICATION NUMBER: US/09/287,849
; PRIOR FILING DATE: 1999-04-07
; PRIOR APPLICATION NUMBER: US 08/818,112
; PRIOR FILING DATE: 1997-03-13
; PRIOR APPLICATION NUMBER: US 08/942,578
; PRIOR FILING DATE: 1997-10-01
; PRIOR APPLICATION NUMBER: US 09/025,197
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 09/056,556
; PRIOR FILING DATE: 1998-04-07
; PRIOR APPLICATION NUMBER: US 09/223,040
; PRIOR FILING DATE: 1998-12-30
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 331
; TYPE: PRY
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:tr1-fusion
US-10-359-460-4

```

```

Query Match      100.0%; Score 736; DB 4; Length 331;
Best Local Similarity 100.0%; Pred. NO. 3.5e-72;
Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

Qy      1 MATTLPVQRRHRSLPFPESESLFAAFPSPFAGLRPTFDRRLMLBDEMKRGYEVRAELPGV 60
Db      8 MATTLPVQRRHRSLPFPESESLFAAFPSPAGLRPTFDRRLMLBDEMKRGYEVRAELPGV 67

Qy      61 DDDVDVIMWDDGQLTKAERTBCKDPGKSEFPAGSVRVVSLPVGADDEDIDATYDKG 120
Db      68 DDDVDVIMWDDGQLTKAERTBCKDPGGRSEFPAGSVRVVSLPVGADDEDIDATYDKG 127

Qy      121 ILTVSVAVSEGGPTEKHIIQRSTN 144
Db      128 ILTVSVAVSEGGPTEKHIIQRSTN 151

RESULT 5
US-10-359-459-4
; Sequence 4, Application US/10359459
; Publication No. US20040013677A1
; GENERAL INFORMATION:
; APPLICANT: Skeiky, Yasir
; APPLICANT: Alderson, Mark

```

```

; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Alderson, Mark
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens
; TITLE OF INVENTION: and Their Uses
; FILE REFERENCE: 014058-009020US
; CURRENT APPLICATION NUMBER: US/09/287,849
; CURRENT FILING DATE: 1999-04-07
; PRIOR APPLICATION NUMBER: US 08/818,112
; PRIOR FILING DATE: 1997-03-13
; PRIOR APPLICATION NUMBER: US 08/942,578
; PRIOR FILING DATE: 1997-10-01
; PRIOR APPLICATION NUMBER: US 09/025,197
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 09/056,556
; PRIOR FILING DATE: 1998-04-07
; PRIOR APPLICATION NUMBER: US 09/223,040
; PRIOR FILING DATE: 1998-12-30
; NUMBER OF SEQ. ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.1
; SEQ. ID NO 24
; LENGTH: 368
; TYPE: PR1
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:tri-fusion
; US-09-287-849-24

```

Qy	1	MATTLPVQHHPSLTPPESESLPAAPSPAGLRPFEDTRLMLRLDEEMKEGRLVRAELPGV	60
Db	9	MATTLPVQHHPSLTPPESESLPAAPSPAGLRPFEDTRLMLRLDEEMKEGRLVRAELPGV	68
Qy	61	DPDKVDIIVNRGQGLTIAERTBOQDFGRSEPAVGSFVRTSLPVGADBDIDIKATYDKG	120
Db	69	DPDKVDIIVNRGQGLTIAERTBOQDFGRSEPAVGSFVRTSLPVGADBDIDIKATYDKG	128
Qy	121	ILTVSAVASEGKPTKEKHQIRSTN	144
Db	129	ILTVSAVASEGKPTKEKHQIRSTN	152

RESULT 8
US-09-287-849-14

```
Sequence 14, Application US/09287849
Patent No. US20020009459A1
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Dillon, Davin C.
APPLICANT: Alderson, Mark
APPLICANT: Campos-Neto, Antonio
APPLICANT: Corixa Corporation
TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens
TITLE OF INVENTION: and Their Uses
FILE REFERENCE: 014058-009020US
CURRENT APPLICATION NUMBER: US/09/287, 849
PRIOR FILING DATE: 1999-04-07
PRIOR APPLICATION NUMBER: US 08/818,112
PRIOR FILING DATE: 1997-03-13
PRIOR APPLICATION NUMBER: US 08/942,578
PRIOR FILING DATE: 1997-10-01
PRIOR APPLICATION NUMBER: US 09/025,197
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 09/056,556
PRIOR FILING DATE: 1998-04-07
PRIOR APPLICATION NUMBER: US 09/223,040
PRIOR FILING DATE: 1998-12-30
NUMBER OF SEQ ID NOS: 46
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 14
LENGTH: 433
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:tetra-fusion
US-09-287-849-14

Query Match      100.0%; Score 736; DB 3; Length 433;
Best Local Similarity 100.0%; Pred. No.5e-72;
Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MATTLPVQRHPRSLFPEFSELFAAFSPFAGLRPTFTRLRLDEMEKGRYEVRAELPGV 60
Db 9 MATTLPVQRHPRSLFPEFSELFAAFSPFAGLRPTFTRLRLDEMEKGRYEVRAELPGV 68

Qy 61 DPKDQVDIMVRDQGLTIKARTEQKDFGRSEFAYGSFVRTVSLPVGADEDDIKATYDKG 120
Db 69 DPKDQVDIMVRDQGLTIKARTEQKDFGRSEFAYGSFVRTVSLPVGADEDDIKATYDKG 128

Qy 121 ILTVSAVSEKGPTEKHQIRSTN 144
Db 129 ILTVSAVSEKGPTEKHQIRSTN 152

RESULT 9
US-10-359-460-14
Sequence 14, Application US/10359460
Publication No. US2003014791A1
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Dillon, Davin C.
APPLICANT: Alderson, Mark
APPLICANT: Campos-Neto, Antonio
APPLICANT: Corixa Corporation
TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens
TITLE OF INVENTION: and Their Uses
FILE REFERENCE: 014058-009020US
CURRENT APPLICATION NUMBER: US/10/359, 460
CURRENT FILING DATE: 2003-02-05
PRIOR APPLICATION NUMBER: US/09/287, 849
PRIOR FILING DATE: 1999-04-07
PRIOR APPLICATION NUMBER: US 08/818,112
PRIOR FILING DATE: 1997-03-13
PRIOR APPLICATION NUMBER: US 08/942,578
PRIOR FILING DATE: 1997-10-01
```

```
PRIOR APPLICATION NUMBER: US 09/025,197
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 09/056,556
PRIOR FILING DATE: 1998-04-07
PRIOR APPLICATION NUMBER: US 09/223,040
PRIOR FILING DATE: 1998-12-30
NUMBER OF SEQ ID NOS: 46
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 14
LENGTH: 433
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:tetra-fusion
US-10-359-460-14

Query Match      100.0%; Score 736; DB 4; Length 433;
Best Local Similarity 100.0%; Pred. No.5e-72;
Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MATTLPVQRHPRSLFPEFSELFAAFSPFAGLRPTFTRLRLDEMEKGRYEVRAELPGV 60
Db 9 MATTLPVQRHPRSLFPEFSELFAAFSPFAGLRPTFTRLRLDEMEKGRYEVRAELPGV 68

Qy 61 DPKDQVDIMVRDQGLTIKARTEQKDFGRSEFAYGSFVRTVSLPVGADEDDIKATYDKG 120
Db 69 DPKDQVDIMVRDQGLTIKARTEQKDFGRSEFAYGSFVRTVSLPVGADEDDIKATYDKG 128

Qy 121 ILTVSAVSEKGPTEKHQIRSTN 144
Db 129 ILTVSAVSEKGPTEKHQIRSTN 152

RESULT 10
US-09-287-849-12
Sequence 12, Application US/09287849
Patent No. US20020009459A1
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Dillon, Davin C.
APPLICANT: Alderson, Mark
APPLICANT: Campos-Neto, Antonio
APPLICANT: Corixa Corporation
TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens
TITLE OF INVENTION: and Their Uses
FILE REFERENCE: 014058-009020US
CURRENT APPLICATION NUMBER: US/09/287, 849
CURRENT FILING DATE: 1999-04-07
PRIOR APPLICATION NUMBER: US 08/818,112
PRIOR FILING DATE: 1997-03-13
PRIOR APPLICATION NUMBER: US 08/942,578
PRIOR FILING DATE: 1997-10-01
PRIOR APPLICATION NUMBER: US 09/025,197
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 09/056,556
PRIOR FILING DATE: 1998-04-07
PRIOR APPLICATION NUMBER: US 09/223,040
PRIOR FILING DATE: 1998-12-30
NUMBER OF SEQ ID NOS: 46
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 12
LENGTH: 856
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:penta-fusion
US-09-287-849-12

Query Match      100.0%; Score 736; DB 3; Length 856;
Best Local Similarity 100.0%; Pred. No.1.3e-71;
Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy 1 MATTLPVQRHRSILFPFSESLFAAFSPFAGLRPTFTRLMLRLEBDMKEGRYEVRAELPGV 60
Db 9 MATTLPVQRHRSILFPFSESLFAAFSPFAGLRPTFTRLMLRLEBDMKEGRYEVRAELPGV 68
Qy 61 DDDKDVIDMRDGLTIKAERTBKDPGGRSEFAYGSFVRTVSLPVGADEDDIKATYDKG 120
Db 69 DDDKDVIDMRDGLTIKAERTBKDPGGRSEFAYGSFVRTVSLPVGADEDDIKATYDKG 128
Qy 121 ILTVSAVSESGKPTKHIQIRSTN 144
Db 129 ILTVSAVSESGKPTKHIQIRSTN 152

RESULT 11
US-10-359-460-12
; Sequence 12, Application US/10359460
; Publication No. US20030147911A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Alderson, Mark
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens
; FILE REFERENCE: 014058-009020US
; CURRENT APPLICATION NUMBER: US/10/359,460
; CURRENT FILING DATE: 2003-02-05
; PRIOR APPLICATION NUMBER: US/09/287,849
; PRIOR FILING DATE: 1999-04-07
; PRIOR APPLICATION NUMBER: US 08/818,112
; PRIOR FILING DATE: 1997-03-13
; PRIOR APPLICATION NUMBER: US 08/942,578
; PRIOR FILING DATE: 1997-10-01
; PRIOR APPLICATION NUMBER: US 09/025,197
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 09/056,556
; PRIOR FILING DATE: 1998-04-07
; PRIOR APPLICATION NUMBER: US 09/223,040
; PRIOR FILING DATE: 1998-12-30
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 856
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:penta-fusion
US-10-359-460-12

Query Match 100.0%; Score 736; DB 4; Length 856;
Best Local Similarity 100.0%; Pred. No. 1.3e-71;
Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MATTLPVQRHRSILFPFSESLFAAFSPFAGLRPTFTRLMLRLEBDMKEGRYEVRAELPGV 60
Db 9 MATTLPVQRHRSILFPFSESLFAAFSPFAGLRPTFTRLMLRLEBDMKEGRYEVRAELPGV 68
Qy 61 DDDKDVIDMRDGLTIKAERTBKDPGGRSEFAYGSFVRTVSLPVGADEDDIKATYDKG 120
Db 69 DDDKDVIDMRDGLTIKAERTBKDPGGRSEFAYGSFVRTVSLPVGADEDDIKATYDKG 128
Qy 121 ILTVSAVSESGKPTKHIQIRSTN 144
Db 129 ILTVSAVSESGKPTKHIQIRSTN 152

RESULT 12
US-10-369-983-13
; Sequence 13, Application US/10369983
; Publication No. US20030235593A1
; GENERAL INFORMATION:

APPLICANT: Skeiky, Yasir
APPLICANT: Guiderian, Jeff
APPLICANT: Reed, Steven
APPLICANT: Corixa Corporation
TITLE OF INVENTION: Fusion Proteins of Mycobacterium Tuberculosis
FILE REFERENCE: 014058-009081US
CURRENT APPLICATION NUMBER: US/10/369,983
CURRENT FILING DATE: 2003-02-18
PRIOR APPLICATION NUMBER: US 60/357,351
PRIOR FILING DATE: 2002-02-15
NUMBER OF SEQ ID NOS: 22
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 13
LENGTH: 875
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:fusion protein
US-10-369-983-13

Query Match 100.0%; Score 736; DB 4; Length 875;
Best Local Similarity 100.0%; Pred. No. 1.3e-71;
Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MATTLPVQRHRSILFPFSESLFAAFSPFAGLRPTFTRLMLRLEBDMKEGRYEVRAELPGV 60
Db 732 MATTLPVQRHRSILFPFSESLFAAFSPFAGLRPTFTRLMLRLEBDMKEGRYEVRAELPGV 791
Qy 61 DDDKDVIDMRDGLTIKAERTBKDPGGRSEFAYGSFVRTVSLPVGADEDDIKATYDKG 120
Db 792 DDDKDVIDMRDGLTIKAERTBKDPGGRSEFAYGSFVRTVSLPVGADEDDIKATYDKG 851
Qy 121 ILTVSAVSESGKPTKHIQIRSTN 144
Db 852 ILTVSAVSESGKPTKHIQIRSTN 875

RESULT 13
US-09-953-510-18
; Sequence 18, Application US/09953510
; Patent No. US20020131975A1
; GENERAL INFORMATION:
; APPLICANT: Horwitz, Marcus A.
; TITLE OF INVENTION: Abundant Extracellular
; Products and Methods for Their Production and Use
; NUMBER OF SEQUENCES: 91
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Kurt A. Maclean
; STREET: 2029 Century Park East, Suite 3800
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90067
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0,
; Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/953,510
; FILING DATE: 14-Sep-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/447,398
; FILING DATE: 23-MAY-1995
; APPLICATION NUMBER: US 08/289,667
; FILING DATE: 12-AUG-1994
; APPLICATION NUMBER: US 08/156,358
; FILING DATE: 23-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Maclean, Kurt A.

REGISTRATION NUMBER: 31,118
REFERENCE/DOCKET NUMBER: 112-272
TELECOMMUNICATION INFORMATION:
TELEPHONE: (310) 788-5000
TELEFAX: (310) 277-1297
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 48 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Mycobacterium tuberculosis
STRAIN: Erdman
SEQUENCE DESCRIPTION: SEQ ID NO: 18:
US-09-953-510-18

Query Match 33.3%; Score 245; DB 3; Length 48;
Best Local Similarity 100.0%; Pred. No. 2.5e-19;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 35 FDTLRMLRLEDEMKRGYEVRAELPGVDPDQVDIMVRDGLTIKAERT 82
Db 1 FDTLRMLRLEDEMKRGYEVRAELPGVDPDQVDIMVRDGLTIKAERT 48

RESULT 14
US-09-953-413-18
Sequence 18, Application US/09953413
Publication No. US20040018209A1
GENERAL INFORMATION:
APPLICANT: Horwitz, Marcus A.
TITLE OF INVENTION: Abundant Extracellular
Products and Methods for Their Production and Use
NUMBER OF SEQUENCES: 91
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kurt A. Maclean
STREET: 2029 Century Park East, Suite 3800
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90067
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0,
Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/953,413
FILING DATE: 14-Sep-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/447,398
FILING DATE: 23-MAY-1995
APPLICATION NUMBER: US 08/289,667
FILING DATE: 12-AUG-1994
APPLICATION NUMBER: US 08/156,358
FILING DATE: 23-NOV-1993
ATTORNEY/AGENT INFORMATION:
NAME: Maclean, Kurt A.
REGISTRATION NUMBER: 31,118
REFERENCE/DOCKET NUMBER: 112-272
TELECOMMUNICATION INFORMATION:
TELEPHONE: (310) 788-5000
TELEFAX: (310) 277-1297
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 48 amino acids

TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Mycobacterium tuberculosis
STRAIN: Erdman
SEQUENCE DESCRIPTION: SEQ ID NO: 18:
US-09-953-413-18

Query Match 33.3%; Score 245; DB 3; Length 48;
Best Local Similarity 100.0%; Pred. No. 2.5e-19;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 35 FDTLRMLRLEDEMKRGYEVRAELPGVDPDQVDIMVRDGLTIKAERT 82
Db 1 FDTLRMLRLEDEMKRGYEVRAELPGVDPDQVDIMVRDGLTIKAERT 48

RESULT 15
US-10-147-255-18
Sequence 18, Application US/10147255
Publication No. US20030152584A1
GENERAL INFORMATION:
APPLICANT: Horwitz, Marcus A.
TITLE OF INVENTION: Abundant Extracellular
Products and Methods for Their Production and Use
NUMBER OF SEQUENCES: 91
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kurt A. Maclean
STREET: 2029 Century Park East, Suite 3800
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90067
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0,
Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/147,255
FILING DATE: 15-May-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/226,539A
FILING DATE: <Unknown>
APPLICATION NUMBER: US 08/447,398
FILING DATE: 23-MAY-1995
APPLICATION NUMBER: US 08/289,667
FILING DATE: 12-AUG-1994
APPLICATION NUMBER: US 08/156,358
FILING DATE: 23-NOV-1993
ATTORNEY/AGENT INFORMATION:
NAME: Maclean, Kurt A.
REGISTRATION NUMBER: 31,118
REFERENCE/DOCKET NUMBER: 112-272
TELECOMMUNICATION INFORMATION:
TELEPHONE: (310) 788-5000
TELEFAX: (310) 277-1297
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 48 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO


```

; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Mycobacterium tuberculosis
; STRAIN: Erdman
; SEQUENCE DESCRIPTION: SEQ ID NO: 18:
US-10-147-255-18
    
```

```

Query Match      33.3%; Score 245; DB 4; Length 48;
Best Local Similarity 100.0%; Pred. No. 2.5e-19;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
    
```

```

CY      35  FDTRLMRLEDEMKGRYEVRAELPGVDPDKVDIMVRDGLTIKAERT 82
          |||||||
Db      1  FDTRLMRLEDEMKGRYEVRAELPGVDPDKVDIMVRDGLTIKAERT 48
    
```

```

Search completed: March 23, 2006, 06:28:12
Job time : 31.2823 secs
    
```

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 23, 2006, 06:17:23 ; Search time 3.45564 Seconds
(without alignments)
1192.746 Million cell updates/sec

Title: US-10-617-038-24

Perfect score: 736

Sequence: 1 MATTLPVQRHPRSLPFFSE.....SVAVSEKPTKHIQIRSTN 144

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 169630 seqs, 2862289 residues

Total number of hits satisfying chosen parameters: 169630

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications RA_New:
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US05_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	736	100.0	331	US-11-201-519-4	Sequence 4, Appl1
2	736	100.0	368	US-11-201-519-24	Sequence 24, Appl1
3	736	100.0	433	US-11-201-519-14	Sequence 14, Appl1
4	736	100.0	856	US-11-201-519-12	Sequence 12, Appl1
5	144.5	19.6	150	US-11-087-039-2558	Sequence 2558, Ap
6	132.5	18.0	153	US-11-087-039-7830	Sequence 7830, Ap
7	130	17.7	210	US-11-087-039-11382	Sequence 11382, A
8	129	17.5	149	US-11-087-039-11382	Sequence 11382, A
9	129	17.5	155	US-11-096-568A-21407	Sequence 21407, A
10	127	17.3	162	US-11-096-568A-16708	Sequence 16708, A
11	126.5	17.2	154	US-11-087-039-8558	Sequence 8558, Ap
12	126.5	17.2	154	US-11-087-039-10259	Sequence 10259, A
13	126.5	17.2	161	US-11-087-039-10976	Sequence 10976, A
14	125.5	17.1	161	US-11-087-039-11804	Sequence 11804, A
15	124	16.8	159	US-11-096-568A-7013	Sequence 7013, Ap
16	123.5	16.8	153	US-11-087-039-12190	Sequence 12190, A
17	121.5	16.5	131	US-11-096-568A-21408	Sequence 21408, A
18	119.5	16.2	128	US-11-087-039-11932	Sequence 11932, A
19	118.5	16.1	156	US-11-087-039-5432	Sequence 5432, Ap
20	111.5	15.0	150	US-11-087-039-1077	Sequence 1077, Ap
21	110.5	15.0	146	US-11-087-039-11787	Sequence 11787, A
22	109	14.8	141	US-11-074-176-304	Sequence 304, App
23	109	14.8	147	US-11-074-176-14	Sequence 14, Appl
24	106.5	14.5	133	US-11-087-039-8784	Sequence 8784, Ap
25	104	14.1	161	US-11-096-568A-14759	Sequence 14759, A

26	104	14.1	198	US-11-096-568A-14758	Sequence 14758, A
27	103	14.0	150	US-11-096-568A-6324	Sequence 6324, Ap
28	103	14.0	153	US-11-096-568A-6323	Sequence 6323, Ap
29	103	14.0	154	US-11-096-568A-6322	Sequence 6322, Ap
30	94.5	12.8	159	US-11-096-568A-7626	Sequence 7626, Ap
31	94.5	12.8	187	US-11-096-568A-7825	Sequence 7825, Ap
32	92	12.5	214	US-11-096-568A-9568	Sequence 9568, Ap
33	92	12.5	239	US-11-096-568A-9567	Sequence 9567, Ap
34	91	12.4	147	US-11-096-568A-21035	Sequence 21035, A
35	88	12.0	124	US-11-096-568A-21036	Sequence 21036, A
36	79	10.7	175	US-10-821-234-1543	Sequence 1543, Ap
37	78.5	10.7	297	US-11-096-568A-18592	Sequence 18592, A
38	78	10.6	136	US-11-096-568A-14425	Sequence 14425, A
39	75.5	10.3	205	US-10-878-556A-199	Sequence 199, App
40	74	10.1	210	US-11-010-239-77	Sequence 77, Appl
41	71.5	9.7	242	US-10-467-657-7306	Sequence 7306, Ap
42	71.5	9.7	1401	US-11-096-568A-28512	Sequence 28512, A
43	71	9.6	149	US-11-096-568A-15190	Sequence 15190, A
44	71	9.6	225	US-11-096-568A-15189	Sequence 15189, A
45	70.5	9.6	505	US-11-087-039-2851	Sequence 2851, Ap

ALIGNMENTS

```

RESULT 1
US-11-201-519-4
; Sequence 4, Application US/11201519
; Publication NO. US2006003764A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Alderson, Mark
; APPLICANT: Campos-Neto, Antonio
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens
; FILE REFERENCE: 014058-009020US
; CURRENT APPLICATION NUMBER: US/11/201,519
; CURRENT FILING DATE: 2005-08-10/359,460
; PRIOR APPLICATION NUMBER: US/10/359,460
; PRIOR FILING DATE: 2003-02-05
; PRIOR APPLICATION NUMBER: US/09/287,849
; PRIOR FILING DATE: 1999-04-07
; PRIOR APPLICATION NUMBER: US 08/818,112
; PRIOR FILING DATE: 1997-03-13
; PRIOR APPLICATION NUMBER: US 08/942,578
; PRIOR FILING DATE: 1997-10-01
; PRIOR APPLICATION NUMBER: US 09/025,197
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 09/056,556
; PRIOR FILING DATE: 1998-04-07
; PRIOR APPLICATION NUMBER: US 09/223,040
; PRIOR FILING DATE: 1998-12-30
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 331
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: tri-fusion
US-11-201-519-4

```

Query Match 100.0%; Score 736; DB 7; Length 331;
Best Local Similarity 100.0%; Pred. No. 2.2e-68;
Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATTLPVQRHPRSLPFFSELPFAFSPFAGLRPTDRLMRLEDEMKEGRYEVRAELPGV 60
DB 8 MATTLPVQRHPRSLPFFSELPFAFSPFAGLRPTDRLMRLEDEMKEGRYEVRAELPGV 67

Qy 61 DPKDVIDIWRDQGLTIKARTEKDPDGRSEFAVGSFVRTVSLPVGADEDDIKATYDKG 120
Db 68 DPKDVIDIWRDQGLTIKARTEKDPDGRSEFAVGSFVRTVSLPVGADEDDIKATYDKG 127
Qy 121 ILTVSAVSESGKPTKEKHIOIRSTN 144
Db 128 ILTVSAVSESGKPTKEKHIOIRSTN 151

RESULT 2

US-11-201-519-24
Sequence 24, Application US/11201519
Publication No. US20060034764A1
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Alderson, Mark C.
APPLICANT: Campos-Neto, Antonio
APPLICANT: Corixa Corporation
TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens
FILE REFERENCE: 014058-009020US
CURRENT APPLICATION NUMBER: US/11/201,519
CURRENT FILING DATE: 2005-08-10
PRIOR APPLICATION NUMBER: US/10/359,460
PRIOR FILING DATE: 2003-02-05
PRIOR APPLICATION NUMBER: US/09/287,849
PRIOR FILING DATE: 1999-04-07
PRIOR APPLICATION NUMBER: US 08/818,112
PRIOR FILING DATE: 1997-03-13
PRIOR APPLICATION NUMBER: US 08/942,578
PRIOR FILING DATE: 1997-10-01
PRIOR APPLICATION NUMBER: US 09/025,197
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 09/056,556
PRIOR FILING DATE: 1998-04-07
PRIOR APPLICATION NUMBER: US 09/223,040
PRIOR FILING DATE: 1998-12-30
NUMBER OF SEQ ID NOS: 46
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 24
LENGTH: 368
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: tetra-fusion
US-11-201-519-24

Query Match 100.0%; Score 736; DB 7; Length 368;

Best Local Similarity 100.0%; Pred. No. 2,5e-68;
Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MATTLPVQRHPRSLFPFSESLFAFPSPAGLRPTFTDRLRLKLEDEMEKGRYEVAELPGV 60
Db 9 MATTLPVQRHPRSLFPFSESLFAFPSPAGLRPTFTDRLRLKLEDEMEKGRYEVAELPGV 68
Qy 61 DPKDVIDIWRDQGLTIKARTEKDPDGRSEFAVGSFVRTVSLPVGADEDDIKATYDKG 120
Db 69 DPKDVIDIWRDQGLTIKARTEKDPDGRSEFAVGSFVRTVSLPVGADEDDIKATYDKG 128
Qy 121 ILTVSAVSESGKPTKEKHIOIRSTN 144
Db 129 ILTVSAVSESGKPTKEKHIOIRSTN 152

RESULT 3

US-11-201-519-14
Sequence 14, Application US/11201519
Publication No. US20060034764A1
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Skeiky, Yasir A.W.

APPLICANT: Dillon, Davin C.
APPLICANT: Alderson, Mark
APPLICANT: Campos-Neto, Antonio
APPLICANT: Corixa Corporation
TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens
FILE REFERENCE: 014058-009020US
CURRENT APPLICATION NUMBER: US/11/201,519
CURRENT FILING DATE: 2005-08-10
PRIOR APPLICATION NUMBER: US/10/359,460
PRIOR FILING DATE: 2003-02-05
PRIOR APPLICATION NUMBER: US/09/287,849
PRIOR FILING DATE: 1999-04-07
PRIOR APPLICATION NUMBER: US 08/818,112
PRIOR FILING DATE: 1997-03-13
PRIOR APPLICATION NUMBER: US 08/942,578
PRIOR FILING DATE: 1997-10-01
PRIOR APPLICATION NUMBER: US 09/025,197
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 09/056,556
PRIOR FILING DATE: 1998-04-07
PRIOR APPLICATION NUMBER: US 09/223,040
PRIOR FILING DATE: 1998-12-30
NUMBER OF SEQ ID NOS: 46
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 14
LENGTH: 433
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: tetra-fusion
US-11-201-519-14

Query Match 100.0%; Score 736; DB 7; Length 433;

Best Local Similarity 100.0%; Pred. No. 3.1e-68;
Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MATTLPVQRHPRSLFPFSESLFAFPSPAGLRPTFTDRLRLKLEDEMEKGRYEVAELPGV 60
Db 9 MATTLPVQRHPRSLFPFSESLFAFPSPAGLRPTFTDRLRLKLEDEMEKGRYEVAELPGV 68
Qy 61 DPKDVIDIWRDQGLTIKARTEKDPDGRSEFAVGSFVRTVSLPVGADEDDIKATYDKG 120
Db 69 DPKDVIDIWRDQGLTIKARTEKDPDGRSEFAVGSFVRTVSLPVGADEDDIKATYDKG 128
Qy 121 ILTVSAVSESGKPTKEKHIOIRSTN 144
Db 129 ILTVSAVSESGKPTKEKHIOIRSTN 152

RESULT 4

US-11-201-519-12
Sequence 12, Application US/11201519
Publication No. US20060034764A1
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Dillon, Davin C.
APPLICANT: Campos-Neto, Antonio
APPLICANT: Corixa Corporation
TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens
FILE REFERENCE: 014058-009020US
CURRENT APPLICATION NUMBER: US/11/201,519
CURRENT FILING DATE: 2005-08-10
PRIOR APPLICATION NUMBER: US/10/359,460
PRIOR FILING DATE: 2003-02-05
PRIOR APPLICATION NUMBER: US/09/287,849
PRIOR FILING DATE: 1999-04-07
PRIOR APPLICATION NUMBER: US 08/818,112
PRIOR FILING DATE: 1997-03-13
PRIOR APPLICATION NUMBER: US 08/942,578

```

; PRIOR FILING DATE: 1997-10-01
; PRIOR APPLICATION NUMBER: US 09/025,197
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 09/056,556
; PRIOR FILING DATE: 1998-04-07
; PRIOR APPLICATION NUMBER: US 09/223,040
; PRIOR FILING DATE: 1998-12-30
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 12
; LENGTH: 856
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: penta-fusion
US-11-201-519-12
```

```

Query Match          100.0%; Score 736; DB 7; Length 856;
Best Local Similarity 100.0%; Pred. No. 7,8e-68;
Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 1 MATTLPVORHPRSLPPEFSELPAAFPSPAGLRPTFTDRLMLRLEDEMKEGRYEVRALPGV 60
    |||||
DB 9 MATTLPVORHPRSLPPEFSELPAAFPSPAGLRPTFTDRLMLRLEDEMKEGRYEVRALPGV 68
QY 61 DDDKVDIMVRDQGLTIKAERTOKDFGRSEFAYGSFVRTVSLPVGADEDDIKATYDKG 120
    |||||
DB 69 DDDKVDIMVRDQGLTIKAERTOKDFGRSEFAYGSFVRTVSLPVGADEDDIKATYDKG 128
QY 121 ILTVSAVASEGKPTTEKHIOIRSTN 144
    |||||
DB 129 ILTVSAVASEGKPTTEKHIOIRSTN 152
```

```

RESULT 5
US-11-087-099-2558
; Sequence 2558, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 2558
; LENGTH: 150
; TYPE: PRT
; ORGANISM: Hordeum vulgare subsp. vulgare
US-11-087-099-2558
```

```

Query Match          19.6%; Score 144.5; DB 7; Length 150;
Best Local Similarity 32.1%; Pred. No. 7,6e-08;
Matches 45; Conservative 27; Mismatches 53; Indels 15; Gaps 5;
```

```

QY 15 PPEFSELPFA-----AFSPAGLRPTFTDRLMLRLEDEMKEGRYEVRALPGVDPDDKV 66
    |||||
DB 10 FDPFADLWADPDTFRSIVPAFSGNSSETAFANAAVDWKETPEAHVFRADLPGVKKE-EV 68
QY 67 DIMVRDQGLTIKAERTOKDFG---RSEFAYGSFVRTVSLPVGADEDDIKATYDKGI 121
    |||||
DB 69 KVEVEDGNVLVYSGERTKEKDKMKHVRVERSSGKFVFRFLPEDAKVEEVKAGLENGV 128
QY 122 LTVSAVASE-GKPTTEKHIOI 140
    |||||
DB 129 LTVTVPKTEVKKPEVKAIEI 148
```

```

RESULT 6
US-11-087-099-7830
; Sequence 7830, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
```

```

; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 7830
; LENGTH: 153
; TYPE: PRT
; ORGANISM: Funaria hygrometrica
US-11-087-099-7830
```

```

Query Match          18.0%; Score 132.5; DB 7; Length 153;
Best Local Similarity 38.7%; Pred. No. 1,4e-06;
Matches 36; Conservative 18; Mismatches 32; Indels 7; Gaps 4;
```

```

QY 54 RAELGVDPDDVDIMVRDQGLTIKAERTOKDFGRSEFAYGSFVRTVSLPVGA 108
    |||||
DB 61 KADLPGLTKE-EVKYVLEGRTLEICGRKKEEVQKSDTHMRERAQGSFMRRLPBGCT 119
QY 109 DEDDIKATYDKGILTVSV-AVSEKPTTEKHIOI 140
    |||||
DB 120 NTDDYKAQVQDGVLTIVTPKQKPKQVRQIEI 152
```

```

RESULT 7
US-11-087-099-11233
; Sequence 11233, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 11233
; LENGTH: 210
; TYPE: PRT
; ORGANISM: Papaver somniferum
US-11-087-099-11233
```

```

Query Match          17.7%; Score 130; DB 7; Length 210;
Best Local Similarity 30.5%; Pred. No. 3,8e-06;
Matches 46; Conservative 23; Mismatches 58; Indels 24; Gaps 6;
```

```

QY 13 SLPEFS-----ELPAFPSPAG-----LRPTFTDRLMLRLEDEMKEGRYEVRAE 56
    |||||
DB 14 NVDFPFLDIWDPOGFPPSTGALTANWQGSSTARETSQLANTRIDWKETPEAHVFRAD 73
QY 57 LPGVDPDDVDIMVRDQGLTIKAERT-----TEOKDFGRSEFAYGSFVRTVSLPVGADE 110
    |||||
DB 74 LPGVTKE-EVKYVEBEGVLDIISGERRRESEKDKMKHVRVERSSGKFLRFLRPENTKM 132
QY 111 DDIKATYDKGILTVSV-AVSEKPTTEKHIOI 140
    |||||
DB 133 DEKATMENGVLTIVCPKVEQRREBVSIEI 163
```

```

RESULT 8
US-11-087-099-11382
; Sequence 11382, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 11382
; LENGTH: 149
; TYPE: PRT
```

```

; ORGANISM: Oryza sativa (japonica cultivar-group)
US-11-087-099-11382
Query Match      17.5%; Score 129; DB 7; Length 149;
Best Local Similarity 33.3%; Pred. No. 3e-06;
Matches 45; Conservative 24; Mismatches 48; Indels 18; Gaps 7;

Qy      12 RSLPFSEFLFAFPSPAGLRPTFDTRLRLDEMKEGRYEVRALPGVDKDVIMWR 71
Db      25 RSLVPATSDRDYTA--AFANARVDW-----KETPES-HVFKADLPQKKE-EVKVEVE 72

Qy      72 DGO-LTIKAERTQKDFG-----RSEFAYGSFVRTVSLPVGADDDIKATYDKGILITVSV 126
Db      73 EGNVLVISGQSKKEDKNDKMHVERSSGQFMRRFLPENAKVDQVAKSMENGVLITVT 132

Qy      127 AVSE-GKPTKHIQI 140
Db      133 PKAEVKKEPVAKIEI 147

RESULT 9
US-11-096-568A-21407
; Sequence 21407, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 21407
; LENGTH: 155
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; NAME/KEY: misc_feature
; LOCATION: (1)..(155)
; OTHER INFORMATION: Cerees Seq. ID no. 12403861
US-11-096-568A-21407
Query Match      17.5%; Score 129; DB 7; Length 155;
Best Local Similarity 29.8%; Pred. No. 3.2e-06;
Matches 45; Conservative 27; Mismatches 61; Indels 18; Gaps 6;

Qy      7 VQHRPSLFPFES-ELFAAPSPAGLRP-----TFDTRLRLDEMKEGRYEVR 55
Db      4 VSRASSNVDFPSLDLMDPDMFRSIVPSAASSGGSETAFANARVDKMETPEAHYFKA 63

Qy      56 ELPGVDKDVIMWRDGOQTLT---KAERTQKDFD--GRSEFAYGSFVRTVSLPVGADE 110
Db      64 DLPEIKKE-EVKVEVEDGNVLVISGKSRSEEDKNDKMHVERSSGQFMRRFLPENAKV 122

Qy      111 DDIAATYDKGILITVSAVSE-GKPTKHIQI 140
Db      123 DEVKAGLENGVLITVTPKTEVKKEPVAKIEI 153

RESULT 10
US-11-096-568A-16708
; Sequence 16708, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 16708
; LENGTH: 162

```

```

; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; NAME/KEY: misc_feature
; LOCATION: (1)..(162)
; OTHER INFORMATION: Cerees Seq. ID no. 12353810
US-11-096-568A-16708
Query Match      17.3%; Score 127; DB 7; Length 162;
Best Local Similarity 31.2%; Pred. No. 5.4e-06;
Matches 39; Conservative 25; Mismatches 51; Indels 10; Gaps 5;

Qy      25 PPSFAGLRPTFDTRLM---RLDEMKEGRYEVRALPGVDKDVIMWRDGO-LTIKAE 80
Db      37 PPSFGGTTTSSSEITAFAGARVDWKETPEAHVFKTVGVGLKKE-EVKVELEDGNVLQISGE 95

Qy      81 RT-----EQKDPDGRSEFAYGSFVRTVSLPVGADDDIKATYDKGILITVSAVSEGPTE- 135
Db      96 RSKQEERKDTWHRVRESGKFLRRFLPENARTBOISAMENGVLITVTPKEAKKADV 155

Qy      136 KHIQI 140
Db      156 KSIQI 160

RESULT 11
US-11-087-099-8558
; Sequence 8558, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 8558
; LENGTH: 154
; TYPE: PRT
; ORGANISM: Quercus suber
US-11-087-099-8558
Query Match      17.2%; Score 126.5; DB 7; Length 154;
Best Local Similarity 31.9%; Pred. No. 5.7e-06;
Matches 46; Conservative 24; Mismatches 55; Indels 19; Gaps 7;

Qy      13 SLPFES-----ELFAAPSPAGLRPT-----FDTRLRLDEMKEGRYEVRALPGVD 62
Db      12 NVDPFESLDLMDPREGSAVAVPSARETTAFT--ARIDKMETPEAHIFKADLPOLKK 69

Qy      63 KDVIMWRDGO-LTIKAERT---EQKDPDGRSEFAYGSFVRTVSLPVGADDDIKATY 117
Db      70 E-EVKVEVEDGNVLQISGERSKEHEEKNDKMHVERSCGFMRRFLPENAKVDQVAKNM 128

Qy      118 DKGILITVSAVSE-GKPTKHIQI 140
Db      129 ENGVLITVMPKEBKPKPAVAKIEI 152

RESULT 12
US-11-087-099-10259
; Sequence 10259, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 10259
; LENGTH: 154
; TYPE: PRT

```


This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 23, 2006, 04:56:29 ; Search time 88.0019 Seconds
(without alignments)
1652.628 Million cell updates/sec

Title: US-10-617-038-25
Perfect score: 1718
Sequence: 1 MPDWTVTYVKSAVQLACR.....PATPRPRIDEVHFVAKDHR 331

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues
Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1718	100.0	331	8	AD137304 M. tuberc
2	915	53.3	344	8	AD137315 M. tuberc
3	529	30.8	104	2	AD137304 M. tuberc
4	529	30.8	104	2	AD137304 M. tuberc
5	417	24.3	332	8	AD137319 M. tuberc
6	106.5	6.2	413	3	AD137319 M. tuberc
7	106.5	6.2	413	3	AD137319 M. tuberc
8	104	6.1	380	6	AD137319 M. tuberc
9	100.5	5.8	1176	4	AD137319 M. tuberc
10	100.5	5.8	1176	4	AD137319 M. tuberc
11	100.5	5.8	1176	4	AD137319 M. tuberc
12	99	5.7	459	4	AD137319 M. tuberc
13	97.5	5.7	459	4	AD137319 M. tuberc
14	97.5	5.7	459	4	AD137319 M. tuberc
15	97	5.6	585	5	AD137319 M. tuberc
16	97	5.6	904	6	AD137319 M. tuberc
17	96.5	5.6	423	6	AD137319 M. tuberc
18	95.5	5.6	3956	8	AD137319 M. tuberc
19	95	5.5	543	7	AD137319 M. tuberc
20	95	5.5	543	6	AD137319 M. tuberc
21	95	5.5	543	9	AD137319 M. tuberc
22	95	5.5	585	8	AD137319 M. tuberc
23	95	5.5	585	8	AD137319 M. tuberc
24	94	5.5	378	6	AD137319 M. tuberc

25	94	5.5	437	7	ABM86609	Abm86609 Rice abio
26	93.5	5.4	188	3	AAy81711	AAy81711 Streptococ
27	93.5	5.4	201	6	ABU00998	ABU00998 S. pneumo
28	93.5	5.4	201	6	ADM92111	ADM92111 S. pneumo
29	93	5.4	431	7	ABM85743	ABM85743 Human pro
30	93	5.4	932	6	ABU34318	ABU34318 Protein e
31	92.5	5.4	519	7	ABO82069	ABO82069 Pseudomon
32	92.5	5.4	737	8	ADJ93777	ADJ93777 AChE1 pro
33	92.5	5.4	7349	6	ABU11385	ABU11385 Protein e
34	92	5.4	200	5	ABP26703	ABP26703 Streptoco
35	92	5.4	200	8	ADV88610	ADV88610 Streptoco
36	92	5.4	200	8	ADV79863	ADV79863 Streptoco
37	92	5.4	200	8	ADV82013	ADV82013 Streptoco
38	92	5.4	333	5	ABP66056	ABP66056 Bifidobac
39	91.5	5.3	750	6	ABU36645	ABU36645 Protein e
40	91.5	5.3	750	6	ABU34712	ABU34712 Protein e
41	91.5	5.3	1346	2	AAW31346	AAW31346 Rat tumou
42	90.5	5.3	457	6	ABU35824	ABU35824 Protein e
43	90.5	5.3	664	7	ABO68223	ABO68223 Pseudomon
44	90.5	5.3	737	8	ADJ93781	ADJ93781 AChE1 pro
45	90.5	5.3	841	6	ABU26085	ABU26085 Protein e

ALIGNMENTS

RESULT 1
AD137304 standard; protein; 331 AA.
AD137304;
22-APR-2004 (first entry)
M. tuberculosis low oxygen induced antigen Rv2032 SEQ ID NO:25.
mycobacterial infection; vaccine; tuberculosis;
Mycobacterium tuberculosis; immunisation; antibacterial; gene therapy;
low oxygen induced antigen.
Mycobacterium tuberculosis.
WO2004006952-A2.
22-JAN-2004.
08-JUL-2003; 2003WO-DK000477.
13-JUL-2002; 2002DK-0001098.
(STAT-) STATENS SERUM INST.
Andersen P, Rosenkrands I, Stryhn A;
WPI; 2004-122778/12.
N-PSDB; AD137349.
Use of one or more polypeptides or their fragments, which are expressed during the latent stage of the mycobacterial infection, and/or nucleic acids encoding the polypeptides, for a therapeutic vaccine against tuberculosis.
Claim 3; SEQ ID NO 25; 76pp; English.
The present invention describes polypeptides or their fragments, which are expressed during the latent stage of a mycobacterial infection, and/or nucleic acids encoding the polypeptides, which are useful for creating a therapeutic vaccine against tuberculosis. Also described: (1) a therapeutic vaccine against tuberculosis comprising one or more polypeptides; (2) a method for creating an animal, including a human being, with tuberculosis caused by virulent mycobacteria, e.g. by Mycobacterium tuberculosis, M. africanum or M. bovis; (3) a method for immunising an animal, including a human being, against tuberculosis

CC caused by virulent mycobacteria; (4) a method of diagnosing tuberculosis
 CC caused by virulent mycobacteria in an animal, including a human being;
 CC (5) a method for diagnosing previous or ongoing infection with a virulent
 CC mycobacterium; and (6) a method of diagnosing Mycobacterium tuberculosis
 CC infection in a subject. The polypeptides have antibacterial activities,
 CC and can be used in vaccines and in gene therapy. The polypeptides are
 CC useful for the manufacture of a therapeutic vaccine for treating an
 CC individual who is infected by a virulent mycobacterium, e.g. M.
 CC tuberculosis, and who is not vaccinated with BCG against tuberculosis.
 CC The present sequence represents a low oxygen induced antigen, which is
 CC used in the exemplification of the present invention.

XX Sequence 331 AA;

Query Match 100.0%; Score 1718; DB 8; Length 331;
 Best Local Similarity 100.0%; Pred. No. 4,1e-172; Mismatches 0; Gaps 0;
 Matches 331; Conservative 0; Indels 0; Gaps 0;

QY 1 MPDTMTTVDVKSANVOLACRAPSLHNSQPMRWIAEDHTVALFLDKRVLVATDHSGRAL 60
 DB 1 MPDTMTTVDVKSANVOLACRAPSLHNSQPMRWIAEDHTVALFLDKRVLVATDHSGRAL 60
 QY 61 LGGCAVLDHFRVAAAAGTTANVERFPNPNPLASIDFSPADFTGHRRLADAILLR 120
 DB 61 LGGCAVLDHFRVAAAAGTTANVERFPNPNPLASIDFSPADFTGHRRLADAILLR 120
 QY 121 RTDRLPFAEPPDMDLVESQRLRTTADTVADTVADIMRPLAASKLTESLRLYDSYH 180
 DB 121 RTDRLPFAEPPDMDLVESQRLRTTADTVADTVADIMRPLAASKLTESLRLYDSYH 180
 QY 121 RTDRLPFAEPPDMDLVESQRLRTTADTVADTVADIMRPLAASKLTESLRLYDSYH 180
 DB 121 RTDRLPFAEPPDMDLVESQRLRTTADTVADTVADIMRPLAASKLTESLRLYDSYH 180
 QY 181 AELFWMTGAFETSGIPHSLSVSAESDRTVFGDFPVVANTDRRPFEGHRSKVLVLT 240
 DB 181 AELFWMTGAFETSGIPHSLSVSAESDRTVFGDFPVVANTDRRPFEGHRSKVLVLT 240
 QY 241 YDNERASLLCGEMLSAVLDTMAAGLATCTLTHITELHASRDVLAALIGPATPOLVR 300
 DB 241 YDNERASLLCGEMLSAVLDTMAAGLATCTLTHITELHASRDVLAALIGPATPOLVR 300
 QY 301 VGLAPEMEPPPATPRRPIDEVFHVRKDR 331
 DB 301 VGLAPEMEPPPATPRRPIDEVFHVRKDR 331

RESULT 2
 ADI37315
 ID ADI37315 standard; protein; 344 AA.

XX AC ADI37315;
 XX DT 22-APR-2004 (first entry)
 XX DE M. tuberculosis low oxygen induced antigen Rv3127 SEQ ID NO:36.
 XX KW Mycobacterial infection; vaccine; tuberculosis;
 XX KM Mycobacterium tuberculosis; immunisation; antibacterial; gene therapy;
 XX OS low oxygen induced antigen.
 XX OS Mycobacterium tuberculosis.
 XX PN WO2004006952-A2.
 XX PD 22-JAN-2004.
 XX PF 08-JUL-2003; 2003WO-DK00477.
 XX PR 13-JUL-2002; 2002DK-00001098.
 XX PA (STAT-) STATENS SERUM INST.
 XX PI Andersen P, Rosenkrands I, Stryhn A;
 XX WPI; 2004-122778/12.
 XX DR N-PSDB; ADI37360.

XX Use of one or more polypeptides or their fragments, which are expressed
 PT during the latent stage of the mycobacterial infection, and/or nucleic
 PT acids encoding the polypeptides, for a therapeutic vaccine against
 PT tuberculosis.

XX Claim 3; SEQ ID NO 36; 76pp; English.

PS The present invention describes polypeptides or their fragments, which
 XX are expressed during the latent stage of a mycobacterial infection,
 CC and/or nucleic acids encoding the polypeptides, which are useful for
 CC creating a therapeutic vaccine against tuberculosis. Also described: (1)
 CC a therapeutic vaccine against tuberculosis comprising one or more
 CC polypeptides; (2) a method for treating an animal, including a human
 CC being, with tuberculosis caused by virulent mycobacteria, e.g. by
 CC Mycobacterium tuberculosis, M. africanum or M. bovis; (3) a method for
 CC immunising an animal, including a human being, against tuberculosis
 CC caused by virulent mycobacteria; (4) a method of diagnosing tuberculosis
 CC caused by virulent mycobacteria in an animal, including a human being;
 CC (5) a method for diagnosing previous or ongoing infection with a virulent
 CC mycobacterium; and (6) a method of diagnosing Mycobacterium tuberculosis
 CC infection in a subject. The polypeptides have antibacterial activities,
 CC and can be used in vaccines and in gene therapy. The polypeptides are
 CC useful for the manufacture of a therapeutic vaccine for treating an
 CC individual who is infected by a virulent mycobacterium, e.g. M.
 CC tuberculosis, and who is not vaccinated with BCG against tuberculosis.
 CC The present sequence represents a low oxygen induced antigen, which is
 CC used in the exemplification of the present invention.

XX Sequence 344 AA;

Query Match 53.3%; Score 915; DB 8; Length 344;
 Best Local Similarity 56.7%; Pred. No. 2.6e-87; Mismatches 182; Conservative 39; Indels 96; Gaps 2;

QY 10 VISAQVLACRAPSLHNSQPMRWIAE---DH-FVALFLDKRVLVATDHSGRALLGCGA 65
 DB 1 VLNRAVLACRAPSVHNSQPMRWVASSGSEHTTVHLFVNHRVTPATDHSGRQALISCGA 60
 QY 66 VLDHFRVAAAAGTTANVERFPNPNPLASIDFSPADFTGHRRLADAILLRDRL 125
 DB 61 VLDHFRVAAAAGTTANVERFPNPNPLASIDFSPADFTGHRRLADAILLRDRL 120
 QY 126 PFAEPPDMDLVESQRLRTTADTVADTVADIMRPLAASKLTESLRLYDSYHAELEW 185
 DB 121 PFDSPMYWHLFEBPALRDVADVKAMLDVSDQRTRLVVASQSEVLRDDPYHALEW 180
 QY 186 WTGAFETSGIPHSLSVSAESDRTVFGDFPVVANTDRRPFEGHRSKVLVLTSDNER 245
 DB 181 WTSPEVLAHGVDPDTLASDAERLVRVDFGRFPVRSYQNRRAELADRSKVLVLTSDNR 240
 QY 246 ASLLRCGEMLSAVLDTMAAGLATCTLTHITELHASRDVLAALIGPATPOLVRGVLAP 305
 DB 241 ADALRCGEVSTLLEBCTMAAGLATCTLTHITELHASRDVLAALIGPATPOLVRGVLAP 300
 QY 306 EMEPPPATPRRPIDEVFHVR 326
 DB 301 PLAAVPAPTRPRPLDSVLQIR 321

RESULT 3
 AAY39048
 ID AAY39048 standard; protein; 104 AA.

XX AC AAY39048;
 XX DT 05-NOV-1999 (first entry)
 XX DE M. tuberculosis recombinant antigen protein L5ER-12.
 XX KW Antigen; diagnosis; detection; infection; antibody; immunisation;
 XX KM vaccine; immunity.

OS Mycobacterium tuberculosis.
XX
XX MO9942118-A2.
XX
XX 26-AUG-1999.
XX
XX 17-FEB-1999; 99WO-US003265.
XX
XX 18-FEB-1998; 98US-00024753.
PR 05-MAY-1998; 98US-00072596.
XX
XX (CORI-) CORIXA CORP.
XX
XX Reed SG, Skeiky YAM, Dillon DC, Campos-Neto A, Houghton R;
PI Vedvick TS, Twardzik DR, Lodes MJ, Hendrickson RC;
XX
XX WPI; 1999-527416/44.
DR N-PSDB; AAZ19173.
XX
XX New polypeptide comprising antigenic portions of M. tuberculosis.
PT
XX Example 3D; Page 268-269; 323pp; English.
PS
XX This invention describes novel recombinant antigens and their encoding
CC nucleic acids derived from Mycobacterium tuberculosis. The novel
CC polypeptides are useful for detecting M. tuberculosis infection in a
CC biological sample by detecting antibodies which bind with the
CC polypeptides, and are useful as vaccines for immunizing against M.
CC tuberculosis infection. The new detection methods are needed as current
CC vaccination strategies do not provide 100% immunity
XX
SQ Sequence 104 AA;
Query Match 30.8%; Score 529; DB 2; Length 104;
Best Local Similarity 100.0%; Pred. No. 2.5e-47;
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 74 MAAAGTTANVERFPNPNDPLHLASIDFSPADFTVTEGHRRLRADAILLRRTRLPFAEPDWM 133
DB 1 MAAAGTTANVERFPNPNDPLHLASIDFSPADFTVTEGHRRLRADAILLRRTRLPFAEPDWM 60
OY 134 DLVESQRLRTVTADTVRIDVIADDMRPDLAASKLTESTLRLYDS 177
DB 61 DLVESQRLRTVTADTVRIDVIADDMRPDLAASKLTESTLRLYDS 104
RESULT 4
AAV39191
ID AAV39191 standard; protein; 104 AA.
XX
XX AAV39191;
AC
XX 05-NOV-1999 (first entry)
DT
XX
XX M. tuberculosis antigen L5ER-12 amino acid sequence.
DE
XX Mycobacterium tuberculosis; M. tuberculosis; antigen; immunogen;
KW immunotherapy; diagnosis; immunisation; vaccine; infection;
KW immune response; skin test.
XX
XX Mycobacterium tuberculosis.
OS
XX MO9942076-A2.
XX
XX 26-AUG-1999.
XX
XX 17-FEB-1999; 99WO-US003268.
XX
XX 18-FEB-1998; 98US-00025197.
PR 05-MAY-1998; 98US-00072967.
XX
XX (CORI-) CORIXA CORP.
XX

PI Reed SG, Skeiky YAM, Dillon DC, Campos-Neto A, Houghton R;
PI Vedvick TS, Twardzik DR, Lodes MJ, Hendrickson RC;
XX
XX WPI; 1999-527409/44.
DR N-PSDB; AAZ19385.
XX
XX New antigens from Mycobacterium tuberculosis useful in diagnostic skin
PT tests and protective or therapeutic vaccines or compositions.
XX
XX Example 3; Page 223; 299pp; English.
PS
XX The present invention describes polypeptides comprising an immunogenic
CC part of a Mycobacterium tuberculosis antigen (Ag). Also described are
CC vaccines and fusion protein containing M. tuberculosis Ag's. M.
CC tuberculosis Ag's, DNAs encoding them, derived fusion proteins and other
CC polypeptides fragments, can be used in pharmaceutical compositions or
CC vaccines to generate a protective or therapeutic immune response to M.
CC tuberculosis and as reagents in skin tests for diagnosis of tuberculosis.
CC Ag can induce proliferation of, or cytokine secretion by, T, B or natural
CC killer cells and/or macrophages in tuberculosis-immune subjects. AAZ19249
CC to AAZ19460 and AAV39083 to AAV39225 are used in the exemplification of
CC the present invention
XX
SQ Sequence 104 AA;
Query Match 30.8%; Score 529; DB 2; Length 104;
Best Local Similarity 100.0%; Pred. No. 2.5e-47;
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 74 MAAAGTTANVERFPNPNDPLHLASIDFSPADFTVTEGHRRLRADAILLRRTRLPFAEPDWM 133
DB 1 MAAAGTTANVERFPNPNDPLHLASIDFSPADFTVTEGHRRLRADAILLRRTRLPFAEPDWM 60
OY 134 DLVESQRLRTVTADTVRIDVIADDMRPDLAASKLTESTLRLYDS 177
DB 61 DLVESQRLRTVTADTVRIDVIADDMRPDLAASKLTESTLRLYDS 104
RESULT 5
AD137319
ID AD137319 standard; protein; 332 AA.
XX
XX AD137319;
AC
XX
XX 22-APR-2004 (first entry)
DT
XX
XX M. tuberculosis low oxygen induced antigen Rv3131 SEQ ID NO:40.
DE
XX mycobacterial infection; vaccine; tuberculosis;
KW Mycobacterium tuberculosis; immunisation; antibacterial; gene therapy;
KW low oxygen induced antigen.
XX
XX Mycobacterium tuberculosis.
OS
XX MO2004006952-A2.
XX
XX 22-JAN-2004.
XX
XX 08-JUL-2003; 2003WO-DK000477.
PR 13-JUL-2002; 2002DK-00001098.
XX
XX (STAT-) STATENS SERUM INST.
XX
XX Andersen P, Rosenkrands I, Stryhn A;
PI WPI; 2004-122778/12.
DR N-PSDB; AD137364.
XX
XX Use of one or more polypeptides or their fragments, which are expressed
PT using the latent stage of the mycobacterial infection, and/or nucleic
PT acids encoding the polypeptides, for a therapeutic vaccine against
PT tuberculosis.

XX Claim 3; SEQ ID NO 40; 76pp; English.
XX
XX The present invention describes polypeptides or their fragments, which
CC are expressed during the latent stage of a mycobacterial infection,
CC and/or nucleic acids encoding the polypeptides, which are useful for
CC creating a therapeutic vaccine against tuberculosis. Also described: (1)
CC a therapeutic vaccine against tuberculosis comprising one or more
CC polypeptides; (2) a method for treating an animal, including a human
CC being, with tuberculosis caused by virulent mycobacteria, e.g. by
CC Mycobacterium tuberculosis, M. africanum or M. bovis; (3) a method for
CC immunising an animal, including a human being, against tuberculosis
CC caused by virulent mycobacteria; (4) a method of diagnosing tuberculosis
CC caused by virulent mycobacteria in an animal, including a human being;
CC (5) a method for diagnosing previous or ongoing infection with a virulent
CC mycobacterium; and (6) a method of diagnosing Mycobacterium tuberculosis
CC infection in a subject. The polypeptides have antibacterial activities,
CC and can be used in vaccines and in gene therapy. The polypeptides are
CC useful for the manufacture of a therapeutic vaccine for treating an
CC individual who is infected by a virulent mycobacterium, e.g. M.
CC tuberculosis, and who is not vaccinated with BCG against tuberculosis.
CC The present sequence represents a low oxygen induced antigen, which is
CC used in the exemplification of the present invention.

XX Sequence 332 AA:

Query Match 24.3%; Score 417; DB 8; Length 332;
Best Local Similarity 36.1%; Pred. No. 9.5e-35;
Matches 120; Conservative 45; Mismatches 139; Indels 28; Gaps 11;

QY 2 PDMWTTDVYKSAVOLACRAPSLNSQPMWIAEDHTVALFLDQRYVATDHSGRRL 61
DB 6 PD---AEVTAVTTLAVRAPSHNTQPMRWVCPTSLSPDMQIRSTDPGRELIL 61
QY 62 GCGAVLDHFRYMAAAGTTAVVERPNNPNDPLHLASIDFSP-----ADFVTEGRLRADA 116
DB 62 SCGVALHCHVALASLQAVKVRFPDKXCHLTIIVQVLVPQADVAL-----AAA 115
QY 117 ILRLRTDLRPPA--EPPDMDLVESQRLTQTVADTVRIADVDMKPELAAASKLTESRL 174
DB 116 IPRRTDRBAYSCWVPFGDIALMAARAAGWMLR-QVSLDNRKAIQAQVADL---HV 171
QY 175 YDSYHAELEFWMTGAFTSEGIPIHSSLSVAESDRVTGGRF--PVAN-TDRRPERGHD 231
DB 172 TDEYLRBELTWSRGYSVAGVPAKNEDPSAPIP-GRLEAGGSLQSPDVLV--ADD 228
QY 232 RSKVLVSTYDNESASLRCGEMLSAVLDDATMGATCTTHTTELASDLV-AALIG 290
DB 229 GAAILAGTETDDRLARLRAGEAASIVLLTATMAGLACCPITEPLEIKTRDAVRAEVFG 288
QY 291 QPATPOLVVGVAPEMEEPPTATPRRIDEV 322
DB 289 AGGYPMILRVGMAPINADPLPPTPRRELQOV 320

RESULT 6
AAB32537
ID AAB32537 standard; protein; 413 AA.

XX AAB32537;
XX
XX 19-JAN-2001 (first entry)
XX
XX S. lavendulae Mmcx encoded protein sequence.
XX
XX Mitomycin; biosynthesis; mitosome ring system; antibiotic; anti-cancer;
XX anti-inflammatory; immune-enhancer; immunosuppressant; asthma;
XX chronic obstructive pulmonary disease; respiratory inflammation;
XX fungicide; pesticide.
XX
XX Streptomyces lavendulae.
XX
XX WO200053737-A2.

XX 14-SEP-2000.
XX
XX 10-MAR-2000; 2000WO-US006394.
XX
XX 12-MAR-1999; 99US-0026965.
XX
XX (MNU) UNIV MINNESOTA.
XX (SHER/) SHERMAN D H.
XX (MAOY/) MAO Y.
XX (VARO/) VAROGLU M.
XX (HEM/) HE M.
XX (SHEL/) SHELDON P C.
XX
XX Sherman DH, Mao Y, Varoglu M, He M, Sheldon PC;
XX WPI; 2000-601980/57.
XX N-PSDB; AAC55838.
XX
XX Novel nucleic acid molecule comprising mitomycin biosynthetic gene
XX cluster useful for cloning mitomycin biosynthetic genes for elucidating
XX the molecular basis of mitosome ring system biosynthesis.
XX
XX Disclosure; Page 390-392; 399pp; English.

XX This invention relates to isolated and purified nucleic acid molecules
XX from the mitomycin biosynthetic gene cluster. Mitomycins are a group of
XX natural products that contain a variety of functional groups, including
XX amino benzquinone and axiridine ring systems. The S. lavendulae
XX mitomycin biosynthetic gene cluster comprises 47 mitomycin genes spanning
XX 55kb of DNA. The invention includes an expression cassette comprising a
XX mitomycin biosynthetic gene operably linked to a promoter, and host cells
XX transformed with the cassette. The nucleotide, and protein sequences and
XX the transformed host cells of the invention result in antileishmanic,
XX antileishmanic, cytotoxic, immunomodulatory, and antibiotic
XX activities. The nucleotide sequences are used to elucidate the molecular
XX basis for the biosynthesis of the mitosome ring system, as well as to
XX engineer the biosynthesis of novel natural products, e.g. antibiotics,
XX anti-inflammatory agents, anti-cancer agents, immune-enhancers,
XX immunosuppressants, agents to treat asthma, chronic obstructive pulmonary
XX disease as well as other disease involving respiratory inflammation, or
XX cholesterol-lowering agents or as crop protection agents (e.g. fungicides
XX or insecticides) as well as biopolymers, e.g. in packaging or biomedical
XX applications, or to engineer PHA monomer synthases. Sequences AAC55782-
XX C55881, AAC55815-C55849 and AAB32485-B32542 represent mitomycin
XX biosynthetic gene cluster DNA sequences and encoded proteins. Sequences
XX AAC55812-C55814, AAC55850-C55856 and AAC55862-C55869 represent PCR
XX primers used in the cloning of the mitomycin biosynthetic genes

XX Sequence 413 AA;

Query Match 6.2%; Score 106.5; DB 3; Length 413;
Best Local Similarity 24.1%; Pred. No. 0.081;
Matches 73; Conservative 32; Mismatches 111; Indels 87; Gaps 14;

QY 2 PDMWTTDVYKSAVOLACRAPSLNSQPMWIAEDHTVALFLDKR-----VIXAT 52
DB 107 PDPRHVADPRLAGVIRQRYPDIRHPLVW-----NELKGFYDEDRRRMDEGYTRLN 160
QY 53 DH-----SGREALLGCGAVLDHFRVMAAAG-----TTAVVERPNNNDP 92
DB 161 VHAELKRRNPRNLVGGYAVVDDHPPAEDAADSRRLRGWGBLQDSADVIRYWN----- 216
QY 93 LHLASIDFSPAD---FVTEGRLRADAILLRTDRLPFAEPMDLVESQLRTTVTADTV 149
DB 217 AHKAGDFFVVVDDSSYRBECHRAIPDE--FAATKE--FAVVTW--VRSTGLPVWMAEW 270
QY 150 RIIDVIADMKP-----ELAAASKLTESRLYDSSYHAELRW-----WTG 188
DB 271 YVEPPAEDDRPGRGDWGBGCHRTAVQATAMMRILAESGASAAFYWNPORTKACPGCLMS 330
QY 189 AFETSEB--IPHSLSVAASDRVTGRRDP-----VVANTDRPFGHRSRVLV 237

Db 331 THLRDGGELPMAGLS-----RFARFPPTGAFRPVAVTCGSGDRVEALADEAAVLV 383

Qy 238 LST 240
::|
384 VNT 386

Db 384 VNT 386

RESULT 7
ADE10309
ID ADE10309 standard; protein; 413 AA.
XX
AC ADE10309;
XX
DT 29-JAN-2004 (first entry)
XX
DE S. Javendulae mitomycin biosynthetic protein MmcX.
XX
KM Mitomycin biosynthetic protein; mitomycin C; antibiotic; MC; apoptosis;
KM tumour hypoxia; cytostatic; anti-tumour agent; cancer.
XX
OS Streptomyces Javendulae.
XX
PN US2003134398-A1.
XX
PD 17-JUN-2003.
XX
PF 12-SEP-2001; 2001US-00953348.
XX
PR 12-SEP-2001; 2001US-00953348.
XX
PA (SHER/) SHERMAN D H.
PA (MAOY/) MAO Y.
PA (VARO/) VAROGLU M.
PA (HEMM/) HE M.
PA (SHEL/) SHELTON P.
XX
PI Sherman DH, Mao Y, Varoglu M, He M, Sheldon P;
XX
XX WPI: 2003-863498/80.
DR N-PSDB; ADE10257.
XX
XX New nucleic acid molecule comprising a sequence having mitomycin
PT biosynthetic gene cluster, useful for enhancing production of
PT antibiotics.
XX
PS Disclosure: SEQ ID NO 139; 308bp; English.
XX
XX The invention relates to an isolated and purified nucleic acid molecule
CC comprising a sequence having mitomycin biosynthetic gene cluster, or its
CC variant or fragment. Also included are an expression cassette comprising
CC the novel nucleic acid molecule (operably linked to a promoter functional
CC in a host cell), a recombinant bacterial host cell in which at least a
CC portion of a nucleic acid molecule comprising mitomycin biosynthetic gene
CC cluster is disrupted (resulting in a recombinant host cell that produces
CC altered levels of mitomycin relative to a corresponding nonrecombinant
CC bacterial host cell), introducing exogenous DNA into a refractory
CC Streptomycin strain, identifying a nucleic acid molecule that is related
CC to at least a portion of a nucleic acid molecule comprising a mitomycin
CC gene cluster, preparing a compound or its salt from the recombinant host
CC cell and a product produced by the recombinant host cell. The nucleic
CC acid encodes a MltR, MltS, MltR, MltQ, MltP, MltN, MltM, MltL,
CC MltK, MltJ, MltI, MltH, MltG, MltF, MltE, MltD, MltC, MltB, MltA and/or
CC MmcA-MmcY. The nucleic acid is useful for enhancing production of
CC mitomycin antibiotics, which induce apoptosis and hence are useful as
CC anti-tumour (via tumour hypoxia) agents and are useful in treating
CC cancer. The gene cluster was isolated from Streptomyces Javendulae. The
CC present sequence is a mitomycin biosynthetic protein of the invention.
XX
XX Sequence 413 AA;

Query Match 6.2%; Score 106.5; DB 7; Length 413;
Best Local Similarity 24.1%; Pred. No. 0.081;
Matches 73; Conservative 32; Mismatches 111; Indels 87; Gaps 14;

Qy 2 PDVTWTTDVIVKSAVQLACRAPSLHNSQPMRWIAEDHTVALFLDKR-----VLVAT 52
Db 107 FDRHYADFAFLAGVIAQRYPDIRHFLWV-----NELKGFYDEDRRRMDYEGYTRLYNL 160
Qy 53 DH-----SGREALLGCACVLDHFRVAMAAAG-----TTANVERPFPNDP 92
Db 161 VHAELKRRNPRLVGGPFAVVDHDPPEADADRREBELGPMGELDQRSADVIRYVN---- 216
Qy 93 LHLASIDFSPAD---FVTEGHRRLRADAILARTDRLPFAEPBPDVLVESQRTTYTADTV 149
Db 217 AHKAGADPVVVDGSSSYTRREGRAIPDE--FAATEK--FADYTRM--VRSVGLPFWMAEW 270
Qy 150 RIDVIADMDRP-----ELAAASKLTESRLYDSSYHAELFW-----WTG 188
Db 271 YVEPFAEDDRPGGRDGMCEGHRTAIVQATAMRLAESGASAFYNNPQRTGKACPGCLWRS 330
Qy 189 AFETISEG--IRHSLVSAESDRYTFGRDFP-----VVANTDRPEEGHRSKTLV 237
Db 331 THLRDGGELPMAGLS-----RFARFPPTGAFRPVAVTCGSGDRVEALADEAAVLV 383

Qy 238 LST 240
::|
384 VNT 386

Db 384 VNT 386

RESULT 8
ABU20272
ID ABU20272 standard; protein; 380 AA.
XX
XX ABU20272;
XX
DT 19-JUN-2003 (first entry)
XX
DE Protein encoded by Prokaryotic essential gene #5799.
XX
KM Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
OS Borrelia cepacia.
XX
PN WO200277183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002WO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
XX WPI: 2003-023926/02.
DR N-PSDB; ACA24142.
XX
XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 25; SEQ ID NO 48196; 1766bp; English.
XX
XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the

CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pcc_sequences

CC Sequence 380 AA;

Query Match 6.1%; Score 104; DB 6; Length 380;

Best Local Similarity 22.2%; Pred. No. 0.13; Mismatches 118; Indels 94; Gaps 15;

Matches 72; Conservative 41; Mismatches 118; Indels 94; Gaps 15;

Qy 11 IKSANOLACRAPSU-----HNSQPRWIAEDHTVALFLDKDRLVYATDHSGRALLGCGA 65

Db 95 VTTGISADRAHTIATAVAHDVREPHIVQGHVFPILAQPGCVLVRAGHT--EA--GCDP 150

Qy 66 VLDHFRMAAAGTANVERPPNNDPLHLASIDSPADPTEGHRRLAADLILRRDRL 125

Db 151 T-----ALAGLT-----PAVVIC--IKDDTMRALPLI 179

Qy 126 PFAEPDMDLVESQLRRTVADTVRIDVADMDPELAAAGKTESLRLYSYHAELFW 185

Db 180 EFAR-----ENHLKGTADLIQYRSRTSIIERIA-----ERTQYTAGTTRAVLY- 226

Qy 186 WTGAFTSEGIPIHSSLVSAESDRVTFGRDPVVAVANTDRPEFGHRSKVLVSTYDNER 245

Db 227 -----RDQPSGSPHIALVRGAPSPDV-----DTPRVV-----HEPLSLV----- 260

Qy 246 ASLLRCGEMLSAVLLDRTAGLACTLTHITELASRLV--VALIGPATQALVRVGLA 304

Db 261 -DLLETGVSTHSWTLDAARDI-----AERDLGVIVLNGCDTKEHLIDVFKA 307

Qy 305 PEMEPPPTPRRPID-EVPHVRAK 328

Db 308 FDEBEKAAALKRRPVDEKTFGIGAQ 332

RESULT 9

AAB70151 ID AAB70151 standard; protein; 1176 AA.

AC AAB70151;

DT 26-APR-2001 (first entry)

DE DNA encoding human synthetase #1.

KW Human; synthetase; immune; inflammation; AIDS; infection; cancer;

KW reproduction.

OS Homo sapiens.

PN WO200107628-A2.

PD 01-FEB-2001.

XX 20-JUL-2000; 2000WO-US019980.

XX 22-JUL-1999; 99US-0144992P.

PR 02-DEC-1999; 99US-0168858P.

XX (INCY-) INCYTE GENOMICS INC.

PI Tang YT, Hillman JL, Bandman O, Yue H, Baughn MR, Lal P, Lu DM;

PI Shah P, Azimzal Y;

XX WPI; 2001-182799/18.

XX New human synthetases, useful for diagnosing, preventing and treating

XX immune disorders, neuronal disorders, reproductive disorders, and cell

XX proliferative disorders such as cancer.

XX Claim 1, Page 89-92; 120pp; English.

XX The present invention relates to human synthetase proteins. These

XX proteins are useful for treating diseases in which they are involved.

XX Such diseases include immune disorders such as inflammation, acquired

XX immunodeficiency syndrome (AIDS), allergies, autoimmune diseases, and

XX infectious cancers, Alzheimer's disease, reproductive disorders and cell

XX proliferative disorders

XX Sequence 1176 AA;

Query Match 5.8%; Score 100.5; DB 4; Length 1176;

Best Local Similarity 22.0%; Pred. No. 1.7;

Matches 57; Conservative 38; Mismatches 87; Indels 77; Gaps 13;

Qy 86 FPNP--NDPPLA-SIDSPADPTEGHRRLADAIL-----LRRTDR 124

Db 50 FPRPYMGRHLHGTETLSKCFRANGYQRLKKGCCFPFGHCTGMPKACADLKEIE 109

Qy 125 LPPAEPDMDLVESQLRRTVADTVRIDVADMDPELAA--ASK--LTESRLYD 176

Db 110 L-VGCPDPFDEBEHEBETSVKTEDIIIKDKAKGKSKAAKAGSSKYQWGMKSLSD 168

Qy 177 S---SYHAELFW-----WTGAFTSEGIPIHSSLVSAES 207

Db 169 BEIVKFSBAEHMWDYFPPPLAIDQLKRMGLKVWRRSPTTIDVNPYDSFVRWQFLTRER 228

Qy 208 DRVTFGRDPVVAVANTDRPEFGHRS-----KVLVSTYDNERASLLRCGE- 253

Db 229 NKIKFGKRYITYPKQCPQCMDDHRCVGBGVQBYTLKLKYLEPPLSLGLK--GRN 286

Qy 254 --MLSAVLDDATYAGLATC 270

Db 287 IFVVAATLRPEITWFGQTCN 305

RESULT 10

AAM79269 ID AAM79269 standard; protein; 1176 AA.

AC AAM79269;

DT 06-NOV-2001 (first entry)

DE Human protein SEQ ID NO 1931.

KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;

KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;

KW tissue growth factor; immunomodulatory; cancer; leukaemia;

KW nervous system disorder; arthritis; inflammation.

OS Homo sapiens.

PN WO200157190-A2.

PD 09-AUG-2001.
 XX
 XX 05-FEB-2001; 2001WO-US004098.
 XX
 XX 03-FEB-2000; 2000US-00496914.
 PR 27-APR-2000; 2000US-00560875.
 PR 20-JUN-2000; 2000US-00588075.
 PR 19-JUL-2000; 2000US-00620325.
 PR 01-SEP-2000; 2000US-00654936.
 PR 15-SEP-2000; 2000US-00653561.
 PR 20-OCT-2000; 2000US-00693325.
 PR 30-NOV-2000; 2000US-00728422.
 XX
 XX (HYSE-) HYSEQ INC.
 XX
 XX Tang YF, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;
 PI Ma Y, Zhao QH, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
 PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
 XX
 XX WPI; 2001-476283/51.
 DR N-PSDB; AAK52402.
 XX
 XX Nucleic acids encoding polypeptides with cytokine-like activities, useful
 PT in diagnosis and gene therapy.
 XX
 XX Claim 20; Page 4333-4335; 6221pp; English.
 XX
 XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
 CC encoded polypeptides (AAM7323-AAM80302) that exhibit activity elating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
 CC (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the
 CC sequence listing were missing at the time of publication
 XX
 XX Sequence 1176 AA;
 SQ
 Query Match 5.8%; Score 100.5; DB 4; Length 1176;
 Best Local Similarity 22.0%; Pred. No. 1.7;
 Matches 57; Conservative 38; Mismatches 87; Indels 77; Gaps 13;
 QY 86 PPNP--NDPLHA-SIDFSPADFTVTEGRLRADAIL-----LRTDR 124
 DB 50 PPYPMMNGRLHGHFTSLSKCEFAVGQRLKGKCLPPFGHCTGMPKACADKLKREIE 109
 QY 125 LPFAEPPDW-DLVESQRLTYTADTVRIDVIADMRPELAA--ASK-----LTESRLYD 176
 DB 110 L-YGCPDPFPDEEBEETSVKTEDIIIKKAKKKKKAAGSSKYQWGMKSLGLSD 168
 QY 177 S---SYHAELEFW-----WTGAFETSEGIPIHS-----SLVSAES 207
 DB 169 EELVFESEAEHMLDYFPPLAIDQLKRMGLKVMRSRSTITDVNPYDSFVRWQFLTLRER 228
 QY 208 DRVTFGRDPFVAVNTDRRPERGHDRS-----KVLVSTYDNERASLLRGE- 253
 DB 229 NKIKGKRYTIVSPKDGQPCMDHROGEGVGPOEYTLTLKLVLEPSPKLSGLK--GKN 286
 QY 254 --MLSAVLLDATMAGLATC 270
 DB 287 IFLVAAATLRPETMFGQTNC 305

RESULT 11
 ID ADL13064 standard; protein; 1176 AA.
 XX
 XX ADL13064;
 AC

XX 06-MAY-2004 (first entry)
 DT
 XX Human steroid-induced C3A liver cell protein #122.
 DE
 XX Hepatotropic; Gene therapy; Wilson disease; liver disorder;
 KM steroid therapy; cirrhosis; hepatitis; human; C3A liver cell.
 XX
 OS Homo sapiens.
 XX US6673549-B1.
 PN
 XX 06-JAN-2004.
 PD
 XX 12-OCT-2001; 2001US-00976594.
 PF
 XX 12-OCT-2000; 2000US-0240409P.
 PR
 XX (INCY-) INCYTE CORP.
 PA
 XX Furness LM, Buchbinder JL;
 PI
 XX WPI; 2004-068610/07.
 DR
 XX Combination useful for preparing a composition for treating liver
 PT disorders associated with steroid therapy, e.g., cirrhosis or hepatitis,
 PT comprises cDNAs that are differentially expressed in response to steroid
 PT treatment.
 XX
 XX Disclosure; SEQ ID NO 793; 141pp; English.
 XX
 XX The invention relates to a combination comprising cDNAs that are
 CC differentially expressed in response to steroid treatment. Also included
 CC are the following: a high throughput method for using a cDNA to detect
 CC differential expression of nucleic acids in a sample; and a high
 CC throughput method of screening molecules or compounds to identify a
 CC ligand that specifically binds a cDNA. The sample is from a subject with
 CC Wilson disease and comparison of a standard defines a stage of that
 CC disease. The high throughput method of screening molecules or compounds
 CC to identify a ligand that specifically binds a cDNA comprises: combining
 CC the combination with molecules or compounds under conditions to allow
 CC specific binding; and detecting specific binding between each cDNA and at
 CC least one molecule or compound. The molecules or compounds are regulatory
 CC proteins. The combination is useful for preparing a composition for
 CC treating liver disorders associated with steroid therapy, e.g., cirrhosis
 CC or hepatitis. The present sequence represents a human protein which is
 CC differentially expressed in steroid-induced C3A liver cells. Note: The
 CC sequence data for this patent did not form part of the printed
 CC specification but was obtained in electronic format directly from USPTO
 CC at seqdata.uspto.gov/sequence.html.
 XX
 XX Sequence 1176 AA;
 SQ
 Query Match 5.8%; Score 100.5; DB 8; Length 1176;
 Best Local Similarity 22.0%; Pred. No. 1.7;
 Matches 57; Conservative 38; Mismatches 87; Indels 77; Gaps 13;
 QY 86 PPNP--NDPLHA-SIDFSPADFTVTEGRLRADAIL-----LRTDR 124
 DB 50 PPYPMMNGRLHGHFTSLSKCEFAVGQRLKGKCLPPFGHCTGMPKACADKLKREIE 109
 QY 125 LPFAEPPDW-DLVESQRLTYTADTVRIDVIADMRPELAA--ASK-----LTESRLYD 176
 DB 110 L-YGCPDPFPDEEBEETSVKTEDIIIKKAKKKKKAAGSSKYQWGMKSLGLSD 168
 QY 177 S---SYHAELEFW-----WTGAFETSEGIPIHS-----SLVSAES 207
 DB 169 EELVFESEAEHMLDYFPPLAIDQLKRMGLKVMRSRSTITDVNPYDSFVRWQFLTLRER 228
 QY 208 DRVTFGRDPFVAVNTDRRPERGHDRS-----KVLVSTYDNERASLLRGE- 253
 DB 229 NKIKGKRYTIVSPKDGQPCMDHROGEGVGPOEYTLTLKLVLEPSPKLSGLK--GKN 286

```

Qy      254 --MLSAVLDDATMAGLATC 270
Db      287 IFLVAATLRPEPTMEQTNC 305

RESULT 12
ADFS4681
ID      ADFS4681 standard; protein; 897 AA.
XX
XX      ADFS4681;
XX
XX      12-FEB-2004 (first entry)
XX
XX      Aconitase protein #2.
XX
XX      Aconitase; acetic acid resistance; vinegar; acetic acid; enzyme.
XX
XX      Gluconacetobacter entanii .
XX
XX      JP2003289867-A.
XX
XX      14-OCT-2003.
XX
XX      01-APR-2002; 2002JP-00098589.
XX
XX      01-APR-2002; 2002JP-00098589.
XX
XX      (MITS-) MITSUKAN GROUP HONSHA KK.
XX
XX      WPI; 2004-038429/04.
XX
XX      N-PSDB; ADFS4680.
XX
XX      Novel protein having aconitase activity, useful for manufacturing vinegar
XX      having high acetic acid concentration.
XX
XX      Claim 2; SEQ ID NO 4; 56pp; Japanese.
XX
XX      The invention relates to an aconitase protein or a modified protein
XX      having one or more amino acid deletions, substitutions or additions and
XX      having aconitase activity. The invention also relates to DNA encoding a
XX      protein of the invention. The DNA is useful for increasing the acetic
XX      acid resistance of microorganisms, preferably acetic acid bacteria
XX      belonging to Acetobacter or Gluconacetobacter genus. A microorganism
XX      containing DNA and having aconitase activity is useful for manufacturing
XX      vinegar. This sequence represents a Gluconacetobacter entanii aconitase
XX      protein of the invention.
XX
XX      Sequence 897 AA;
XX
Query Match      5.8%; Score 99; DB 8; Length 897;
Best Local Similarity 24.1%; Pred. No. 1.6;
Matches 68; Conservative 31; Mismatches 85; Indels 98; Gaps 15;

Qy      46 DRVLVATDHSRGRLLG-----CGAVLDHFRVAAAGTTANVERFRPNNDPLHLASIDF 100
Db      268 DLVLTVTOMLRKKGVGFVFFGPALDHLFVADRA--TINMA-----PEIGATGCF 318

Qy      101 SPADPVT-----EGHRLRADAILRLRTDRLPFAE---PPMDLVESQLRITVTADT 148
Db      319 FPPVDLTLDYLRQGRREHRIKLTAEYLKAGQMPRHASAPRVFTDLEMLLETTPVS-- 376

Qy      149 VRIVVIADDMRPPELLAASKLTESLRLYDSYHAEIFMTWTGAFETSEGIPIHSSLVSAESD 208
Db      377 -----IAGPKRPQDRVV-----LKGADKAFKEKL--TG-----SLGVPE-----ADKD 412

Qy      209 RVTEGRDFPVVANTDRREFGHDSKVLVSTYDNERASLRGCBMLSAVLDDATMAGLA 268
Db      413 KKA-----KAGNTNY--EIGHG-----DVIATIT 435

Qy      269 TCTLTHITELHASRDLVALIGOPA-----TPQALVVRGLAP 305
Db      436 SCTNTS-----NPAVLIAAGLVAKKARALGLPKPKPVKTSIAP 473

```

```

RESULT 13
AAU48145
ID      AAU48145 standard; protein; 459 AA.
XX
XX      AAU48145;
XX
XX      27-FEB-2002 (first entry)
XX
XX      Propionibacterium acnes immunogenic protein #9041.
XX
XX      SAPHO syndrome; synovitis; acne; pustulosis; hyperostosis; osteomyelitis;
XX      uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
XX      inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
XX      dermatological; osteopathic; neuroprotectant.
XX
XX      Propionibacterium acnes.
XX
XX      WO200181581-A2.
XX
XX      01-NOV-2001.
XX
XX      20-APR-2001; 2001WO-US012865.
XX
XX      21-APR-2000; 2000US-0199047P.
XX
XX      02-JUN-2000; 2000US-0208841P.
XX
XX      07-JUL-2000; 2000US-0216747P.
XX
XX      (CORI-) CORIXA CORP.
XX
XX      Skeiky YAM, Persing DH, Mitcham UL, Mang SS, Bhactia A;
XX      L'malomeuve J, Zhang Y, Jen S, Carter D;
XX      WPI; 2001-616774/71.
XX
XX      N-PSDB; AAS59542.
XX
XX      Propionibacterium acnes polypeptides and nucleic acids useful for
XX      vaccinating against and diagnosing infections, especially useful for
XX      treating acne vulgaris.
XX
XX      Example 1; SEQ ID NO 9340; 1069pp; English.
XX
XX      Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
XX      polypeptides. The proteins and their associated DNA sequences are used in
XX      the treatment, prevention and diagnosis of medical conditions caused by
XX      P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
XX      pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.
XX      P. acnes is also involved in infections of bone, joints and the central
XX      nervous system, however it is particularly involved in the inflammatory
XX      lesions associated with acne vulgaris. A method for detecting the
XX      presence or absence of P. acnes in a patient comprises contacting a
XX      sample with a binding agent that binds to the proteins of the invention
XX      and determining the amount of bound protein in the sample. The
XX      polypeptides may be used as antigens in the production of antibodies
XX      specific for P. acnes proteins. These antibodies can be used to
XX      downregulate expression and activity of P. acnes polypeptides and
XX      therefore treat P. acnes infections. The antibodies may also be used as
XX      diagnostic agents for determining P. acnes presence, for example, by
XX      enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
XX      this patent did not form part of the printed specification, but was
XX      obtained in electronic format directly from WIPO at
XX      ftp.wipo.int/pub/published_pct_sequences
XX
XX      Sequence 459 AA;
XX
Query Match      5.7%; Score 97.5; DB 4; Length 459;
Best Local Similarity 21.2%; Pred. No. 0.85;
Matches 73; Conservative 33; Mismatches 129; Indels 109; Gaps 13;

Qy      70 FRVAM-AAAGTTANVERFRPNNDPLHLASIDFSPADPVTGEGH---RLRADAILRLRTDR 124
Db      105 YTVAVDAAHGITTGIDAVQGARARAVLADPSSPSDLVREGHVLPLRARCGVLERRG-- 162

```


QY 125 LPEAPPDWDLVESQLRTTVDTVRI-----DVIAD--MRPELAASKLTES 171
 DB 163 -----HTEAAVDLARIAGLEPVGLIGELVDDGMCRLADADLARTEG 206
 QY 172 LRLYDSSYHAELFWMTGAFETSEGIPIHSSLSVSAESDRVTPGRDPVAVANDR----- 224
 DB 207 LALTTID--QLAQWRQAHDPGPAVPTQRTVLAASAHLPTRHGGFATGYRDNCTGVEHV 263
 QY 225 -----RPEFGHDSKVLVLSTYDNERASLRLCGEMLS-----AVLL 260
 DB 264 LLVEKGIADDDGGPAWVRVHSECLTGDAIGSLRCD--CGDQLAAAOQHVCRRHGAIIL 320
 QY 261 DATMAGLATCTTLTHITELHA-----SRDLVAA----- 287
 DB 321 LRDEGRATGGLINKIAAYKADGGGLDTVDAQTHLGLPVDAREYGAVALIANIGITSVRL 380
 QY 288 LIGQATPQALVRVGLAPMEBPPATRRRPIDVFNHRAKDR 331
 DB 381 LTNNPAKAEALRDGGL--EVTMSPLETSTREPDERY-LRTKRRD 421

RESULT 14

ABM44664 standard; protein; 459 AA.

ABM44664;

20-OCT-2003 (first entry)

Propionibacterium acnes predicted ORF-encoded polypeptide #9340.

Acne vulgaris; antiseborrheic; dermatological; antibacterial;

immunostimulant; immune response; vaccine.

Propionibacterium acnes.

WO2003033515-A1.

24-APR-2003.

11-OCT-2002; 2002WO-US032727.

15-OCT-2001; 2001US-00978825.

(CORI-) CORIXA CORP.

Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL,

Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D,

Barth B, Valliave-Douglas J;

WPI; 2003-381789/36.

N-P8DB; ACF64471.

New Propionibacterium acnes polypeptides and polymucleotides encoding the

polypeptide, useful for diagnosing, preventing or treating acne vulgaris,

or for stimulating an immune response specific for a P. acnes protein.

Example 1; SEQ ID NO 9340; 1481pp; English.

The invention relates to an isolated polymucleotide (ACF64435-ACF64733)

encoding a Propionibacterium acnes protein. The invention also relates to

polymucleotides encoded by the polymucleotides (ABM35624-ABM64536) and to

immunogenic fragments of P. acnes polypeptides. The invention

additionally encompasses expression vectors and host cells comprising a

polymucleotide of the invention; antibodies against polypeptides of the

invention; fusion proteins comprising a polypeptide of the invention; a

method for stimulating an immune response specific for a P. acnes

polypeptide and an isolated T cell population comprising T cells prepared

via this method; a vaccine composition (comprising P. acnes polypeptides,

polymucleotides, antibodies, fusion proteins, T cell populations, or

antigen-presenting cells that express the polypeptide); a method and kit

for detecting or determining the presence or absence of P. acnes in a

patient; and a method for inhibiting the development of P. acnes in a

CC patient. The P. acnes polypeptides, polymucleotides, antibodies, fusion
 CC proteins, T cell populations or antigen-presenting cells that express the
 CC polypeptides are useful for diagnosing, preventing or treating acne
 CC vulgaris, or for stimulating an immune response specific for a P. acnes
 CC protein. The polymucleotides can also be used as probes or primers for
 CC nucleic acid hybridisation. The vaccine composition is useful for the
 CC stimulation of an immune response against P. acnes, or for treating acne,
 CC and the kit is useful for performing a diagnostic assay. The present
 CC sequence represents a polypeptide predicted to be encoded by an ORF (open
 CC reading frame) contained within the P. acnes polymucleotides of the
 CC invention. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from Wipo. int/pub/published_pct_sequences

CC Sequence 459 AA;

Query Match 5.7%; Score 97.5; DB 6; Length 459;
 Best Local Similarity 21.2%; Pred. No. 0.85;
 Matches 73; Conservative 33; Mismatches 129; Indels 109; Gaps 13;

QY 70 FRVAM-AAAGTTAVVERFPNNDPLHLASIDFSPADFVTEGH---RLRADAILRRTRD 124
 DB 105 YTVAVDAAHGTTTGIDAVQRAKRTARVLADBSSTSDLVREHVLPLABAGVLERRG-- 162
 QY 125 LPEAPPDWDLVESQLRTTVDTVRI-----DVIAD--MRPELAASKLTES 171
 DB 163 -----HTEAAVDLARIAGLEPVGLIGELVDDGMCRLADADLARTEG 206
 QY 172 LRLYDSSYHAELFWMTGAFETSEGIPIHSSLSVSAESDRVTPGRDPVAVANDR----- 224
 DB 207 LALTTID--QLAQWRQAHDPGPAVPTQRTVLAASAHLPTRHGGFATGYRDNCTGVEHV 263
 QY 225 -----RPEFGHDSKVLVLSTYDNERASLRLCGEMLS-----AVLL 260
 DB 264 LLVEKGIADDDGGPAWVRVHSECLTGDAIGSLRCD--CGDQLAAAOQHVCRRHGAIIL 320
 QY 261 DATMAGLATCTTLTHITELHA-----SRDLVAA----- 287
 DB 321 LRDEGRATGGLINKIAAYKADGGGLDTVDAQTHLGLPVDAREYGAVALIANIGITSVRL 380
 QY 288 LIGQATPQALVRVGLAPMEBPPATRRRPIDVFNHRAKDR 331
 DB 381 LTNNPAKAEALRDGGL--EVTMSPLETSTREPDERY-LRTKRRD 421

RESULT 15

ABP69805 standard; protein; 585 AA.

ABP69805;

20-JAN-2003 (first entry)

Human polypeptide SEQ ID NO 1852.

Human; genome mapping; gene therapy; food supplement; virus; fungus;

cell-proliferative disorder; neurodegenerative disease; bacterial;

Parkinson's disease; Alzheimer's disease; autoimmune disease; disease;

multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;

arthritis; cytostatic; immunomodulator; nootropic; neuroprotective;

antiparkinsonian; antidiabetic; immunosuppressive; dermatological;

haemostatic; vulnery; fungicide; antibacterial; virucide; protozoacide;

antiarthritic.

Homo sapiens.

WO200270539-A2.

12-SEP-2002.

05-MAR-2002; 2002WO-US005095.

05-MAR-2001; 2001US-00799451.

Result No.	Score	Query	Length	DB	ID	Description
1	1718	100.0	331	2	G70942	hypothetical prote
2	912	53.1	344	2	F70922	hypothetical prote
3	417	24.3	332	2	D70645	hypothetical prote
4	295	17.2	330	2	T36944	hypothetical prote
5	252	14.7	335	2	T37034	hypothetical prote
6	171	10.0	309	2	T37033	hypothetical prote
7	143	8.3	715	2	B70741	probable moey prot
8	121	7.0	372	2	B70661	hypothetical prote
9	109.5	6.4	213	2	A84250	NADH oxidase [impo
10	109	6.3	409	2	C87329	hypothetical prote
11	108	6.3	721	2	A87128	conserved hypothet
12	104.5	6.1	382	2	AG3161	conserved hypothet
13	102.5	6.0	395	2	T34968	redy protein - Str
14	99	5.8	434	2	F75425	rRNA nucleotidyltr
15	98	5.7	373	2	T34743	hypothetical prote
16	96.5	5.6	481	2	C64945	probable rRNA meth
17	96	5.6	2082	2	T37056	probable multi-dom
18	95.5	5.6	481	2	A88947	probable rRNA meth
19	95.5	5.6	481	2	E85795	probable nucleolar
20	94	5.5	602	2	T50974	related to hxb pro
21	94	5.5	1111	2	T38407	leucyl-L-riua synth
22	93.5	5.4	201	2	E95072	nitroreductase fam
23	93.5	5.4	765	2	H84247	adaptive-response
24	92.5	5.4	547	2	G71307	probable oligopept
25	92.5	5.4	1131	2	T15617	hypothetical prote
26	92	5.4	1186	2	T19334	hypothetical prote
27	91.5	5.3	559	2	E84213	hypothetical prote
28	91.5	5.3	750	2	H70711	probable mutb prote
29	91	5.3	516	1	S72937	hypothetical prote

30	91	5.3	517	2	G87032	conserved hypothet
31	91	5.3	635	2	F70874	probable membrane
32	90.5	5.3	383	2	A95871	probable glycosylt
33	90.5	5.3	421	2	S35205	citrate synthase-1
34	90.5	5.3	457	2	G87003	pyruvate oxidoredu
35	90	5.2	881	2	F84404	valyl-L-lysine synthet
36	89.5	5.2	201	2	B97940	nitroreductase (EC
37	89.5	5.2	1173	2	T25539	hypothetical prote
38	89.5	5.2	1615	2	B49502	protein-tyrosine-p
39	89.5	5.2	1767	2	A49502	protein-tyrosine-p
40	89.5	5.2	1802	2	H88444	protein C26E6.12 l
41	89	5.2	283	2	C94321	hypothetical prote
42	89	5.2	448	2	G70977	hypothetical prote
43	89	5.2	1143	2	A69465	DNA-directed DNA p
44	89	5.2	1150	2	S58775	myl protein - smu
45	88.5	5.2	1059	2	T21891	hypothetical prote

ALIGNMENTS

RESULT 1
 G70942
 hypothetical protein Rv2032 - Mycobacterium tuberculosis (strain H37RV)
 C/Species: Mycobacterium tuberculosis
 C/Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #ext_change 09-Jul-2004
 C/Accession: G70942
 R/Cole, S.T.; Broesch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.
 ; Connor, R.M.; Davies, R.; Devlin, K.; Feltham, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
 ; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squire, S.
 ; Nature 393, 537-544, 1998
 A/Authors: Sgares, R.; Suleston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A/Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
 A/Reference number: A70500; PMID:98295987; PMID:9634230
 A/Accession: G70942
 A/Status: preliminary; nucleic acid sequence not shown; translation not shown
 A/Molecule type: DNA
 A/Residues: 1-331 <COL>
 A/Cross-references: UNIPROT:O53476; UNIPARC:UPI00001652C1; GB:AL021899; GB:AL123456; NID
 A/Experimental source: strain H37RV
 C/Genetics:
 A/Gene: Rv2032
 C/Superfamily: Mycobacterium hypothetical protein Rv3131

Query Match	100.0%	Score 1718;	DB 2;	Length 331;
Best Local Similarity	100.0%	Pred. No. 4.4e-142;		
Matches 331; Conservative	0;	Mismatches	0;	Gaps 0;

[illegible]

RESULT 2

F70922
hypothetical protein Rv3127 - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C:Accession: F70922
R: Cole, S.T.; Broese, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Felwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A: Authors: Squares, R.; Sultson, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A: Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A: Reference number: A70500; PMID:9829587; PMID:9634230
A: Accession: F70922
A: Status: preliminary; nucleic acid sequence not shown; translation not shown
A: Molecule type: DNA
A: Residues: 1-344 <COL>
A: Cross-references: UNIPROT: O05600; UNIPARC: UP1000016533C; GB: Z95150; GB: AL123456; NID: 9
A: Experimental source: strain H37RV
C: Gene: Rv3127
C: Superfamily: Mycobacterium hypothetical protein Rv3131

Query Match 53.1%; Score 912; DB 2; Length 344;
Best Local Similarity 56.4%; Pred. No. 8.9e-72;
Matches 181; Conservative 40; Mismatches 96; Indels 4; Gaps 2;
QY 10 VIKSAVOLACRAPSLHNSQPMRWIAE--DH-TVALPLDKDRLVATDHSGREALLGCGA 65
1 MLKNAVLACRAPSHNSQPMRWIAE--DH-TVALPLDKDRLVATDHSGREALLGCGA 60
DB 66 VLDHFRVMAAAGTTANVERFPNPDLHLASIDFSPADVTGHRRLADAILLRTRDL 125
61 VLDHFRVMAAAGTTANVERFPNPDLHLASIDFSPADVTGHRRLADAILLRTRDL 120
QY 126 PFAEPPDMDLVESQRLTVDYTRIDYADDMRPELAASKLTESRLVYSSYHAELEW 185
121 PFAEPPDMDLVESQRLTVDYTRIDYADDMRPELAASKLTESRLVYSSYHAELEW 180
DB 186 WTGAFTSEGIPIHSSILVSAESDRVTFGRDPVAVANTDRPEFGHRSKVLVLTSDNER 245
181 WTGAFTSEGIPIHSSILVSAESDRVTFGRDPVAVANTDRPEFGHRSKVLVLTSDNER 240
QY 246 ASLIRGEMLSAVLIDATAGLACTTTLHTELHASRDVLAALIGQAPTAQALVRVGLAP 305
241 ASLIRGEMLSAVLIDATAGLACTTTLHTELHASRDVLAALIGQAPTAQALVRVGLAP 300
DB 306 EMEPPATPRPRRIDEVPHVR 326
301 EMEPPATPRPRRIDEVPHVR 321
DB

RESULT 3

D70645
hypothetical protein Rv3131 - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C:Accession: D70645
R: Cole, S.T.; Broese, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Felwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A: Authors: Squares, R.; Sultson, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A: Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A: Reference number: A70500; PMID:9829587; PMID:9634230
A: Accession: D70645
A: Status: preliminary; nucleic acid sequence not shown; translation not shown
A: Molecule type: DNA
A: Residues: 1-332 <COL>
A: Cross-references: UNIPROT: P95195; UNIPARC: UP1000016533D; GB: Z83867; GB: AL123456; NID: 9
A: Experimental source: strain H37RV
C: Gene: Rv3131

C:Superfamily: Mycobacterium hypothetical protein Rv3131

Query Match 24.3%; Score 417; DB 2; Length 332;
Best Local Similarity 36.1%; Pred. No. 1.2e-28;
Matches 120; Conservative 45; Mismatches 139; Indels 28; Gaps 11;
QY 2 PDMWTTDVIKSAVOLACRAPSLHNSQPMRWIAEDHIVALLFLDKDRLVATDHSGRALL 61
6 PD-----AETKRVTLTLVRAPSHINTQPMRWKCPISLEIFSRBDQRLSTDPGRELIL 61
DB 62 GCGAVLDFHFRVMAAAGTTANVERFPNPDLHLASIDFSP-----ADFTGHRRLADA 116
62 SCGVALHCHCVALLSLGQAKVNRFPDPKCRHLATIGVPLVPDQADVAL-----AAA 115
QY 117 ILARTRDLRPEA--EPDMDLVESQRLTVDYTRIDYADDMRPELAASKLTESRL 174
116 IPRRTTRRAVYSCMPVPGGDIALMAAARAGVWLK-QVALDRMKALVAGAVLD--HV 171
DB 175 YDSSYHAELEFWTGAFTSEGIPIHSSILVSAESDRVTFGRDF--PVVAN-TDRRPEFGH 231
172 TDEEYLRLELTWISGRYGVAGVPARNRPPSDPAPIP-GRIFAGPGISQPSDVLK--ADD 228
QY 232 RSKVLVSTYDNERASLRCGEMLSAVLIDATAGLACTTTLHTELHASRDVLAALIG 290
229 GAAILALGTETDRLARLRAGEAASIVLTLRTAMGLACCETPEPLIAKTRDAVRAVFG 288
DB 291 QPATQALVRVGLAPEMEPPATPRRIDEV 322
289 AGVPQMLRVGMARVADPLPPTRRRELQV 320
DB

RESULT 4

T36944
hypothetical protein SCU1.11 - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: T36944
R: Seeger, K.U.; Harris, D.; Thomson, N.R.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, September 1999
A: Reference number: Z21607
A: Accession: T36944
A: Status: preliminary; translated from GB/EMBL/DDBB
A: Molecule type: DNA
A: Residues: 1-330 <SEE>
A: Cross-references: UNIPROT: Q9RJ03; UNIPARC: UP100000DB389; EMBL: AL109962; PIDN: CAB53129.1
A: Experimental source: strain A3 (2)
C: Gene: SCOE1.11
C: Superfamily: Mycobacterium hypothetical protein Rv3131

Query Match 17.2%; Score 295; DB 2; Length 330;
Best Local Similarity 31.0%; Pred. No. 5.4e-18;
Matches 108; Conservative 49; Mismatches 129; Indels 62; Gaps 17;
QY 11 IKAIVOLACRAPSLHNSQPM--RWIADHTVALFLDKDRLVATDHSGREALLGCGAVLD 68
11 VTLTVAEATTAAPSMHNAQPMWRFRFLAERVLVLRADPARAMPSSDGNRLHIGCGAALF 70
DB 69 HFRVMAAAGTTANVERFPNPDLHLASIDFSPADVTGHRRLADAILLRTRDL 125
71 MLKNAVLACRAPSHNSQPMRWIAE--DH-TVALPLDKDRLVATDHSGREALLGCGA 60
QY 126 PFAEPPDMDLVESQRLTVDYTRIDYADDMRPELAASKLTESRL 173
130 PFSER---DIPEN-VATLQAAAARAGATLPLFGMLHVDITLLELVND-----AESRD 177
DB 174 LYSSYHAELEFWT---GAFETIS-EGIPHSSILVSAESDRVTFGRDPVAVANTDRRPEFG 229
178 LVAAEDEDLWTRTRIGGEVDITAVDGP-AVARGPKSKSGKAVRPF-----SGRRPV-- 229
QY 230 HDRSK-----VLVSTYDNERASLRCGEMLSAVLIDATAGLACTTTLHTELHAS 281
230 PDRGSAFERTPTLALGTGGRDMDLAAQQLERVLLEGTURGLSTLSHPLE---D 286
DB

Db 596 TNERV-----SASSALATVYDQATLTDPARGSGANQAVIYAQOGLAVQMSPIFLYAR 651
 Qy 276 -----TELHA-----SRDLVALIIGOPATPOLAVVGLAPMEBPPATPRRP 318
 Db 652 GRHDLDOGSPHPAQLHRLQDPRFLVAP--GKSGHEVLIRLPLFAF-----PFSVCSRR- 704
 Qy 319 IDEVFHYRAKQHR 331
 Db 705 --RVRAHPIPEPDR 715

RESULT 8

B70661
 C:Species: Mycobacterium tuberculosis
 C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
 C:Accession: B70661
 R: Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feldwell, T.; Gentles, S.; Hamlin, N.; Holtz, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Nature 393, 537-544, 1998
 A: Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A: Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
 A: Reference number: A70500; MUID: 98295987; PMID: 9634230
 A: Accession: B70661
 A: Status: preliminary; nucleic acid sequence not shown; translation not shown
 A: Molecule type: DNA
 A: Residues: 1-372 <COL>
 A: Cross-references: UNIPROT: P95233; UNIPARC: UP100000CCBF7; GB: Z83860; GB: AL123456; NID: 5
 A: Experimental source: strain H37RV
 C: Genetics:
 A: Gene: Rv2337c
 C: Superfamily: Mycobacterium hypothetical protein Rv3131

Query Match 7.0%; Score 121; DB 2; Length 372;
 Best Local Similarity 23.9%; Pred. No. 0.0095;
 Matches 86; Conservative 35; Mismatches 121; Indels 124; Gaps 17;

Qy 13 SAVOACRAPSLN-----SOPRWIAEDHTVALFLDKDRVLATDHS 55
 Db 20 SLVEAALAPSDARREYQLEHAGRVLMGDQTR-----SAPENR 61
 Qy 56 GREALLCGAVLDHFRVMAAAGTTANVERPNNDPLHLASIDPSPADFTGHRLLAD 115
 Db 62 RIMSLVAIGAAVENKLRAGLGETKVCWPDGPNGLVAEIP-----VDRLLPQTRVD 115
 Qy 116 ----ALLRRTD-RLPFAEPPDMDLVESQRTVTATDVRIDVI-----ADNRPELAA 164
 Db 116 PIBGAIERRKTRNRRRGPR--LSQGL-GLASAEATGIDGIQLMWFSPETRKQILR 171
 Qy 165 ASKLTESLRVDSSTYHAEF-----WMTGAFETSEGIPIHSSLVSAESDRTVFGDRFP 217
 Db 172 LVRLAEERPRSRRLHELFSAVAFDGLMTS--SDGLPRGSLVEVAM----- 218
 Qy 218 VVANTDRPRF-CHDRSKVTLVSTYDNERASLACGEMLSAVLLDATMAGLATCTLR-HI 275
 Db 219 -----NRPMRGIRHMRVL-----RLRTVGCHNAGI--RAAVLPRLPHV 259
 Qy 276 TELHASRDIVA-----ALIGAPTPQALVAVGLAPMEBPPATPRRP 319
 Db 260 GALTTSIDLISGALTAGAVFERIWLRTTLGAEIOPFAASAVLSLPACENAP----- 312
 Qy 320 DEVFHYRA 327
 Db 313 ----HVR 316

RESULT 9

A84250
 C:Species: Halobacterium sp. NRC-1
 C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
 C:Accession: A84250

R: Ng, W.V.; Kennedy, S.P.; Mahites, G.G.; Bergiste, B.; Pan, M.; Shukla, H.D.; Laaky, S.; Lethausser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablor Jung, K.H.; Alam, M.; Freilias, T.
 Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
 A: Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ehardt, H.; Lowe, T.M.; Li, A.; Title: Genome sequence of Halobacterium species NRC-1.
 A: Reference number: A84160; MUID: 20504483; PMID: 11016950
 A: Accession: A84250
 A: Status: preliminary
 A: Molecule type: DNA
 A: Residues: 1-213 <STO>
 A: Cross-references: UNIPROT: Q9HQZ4; UNIPARC: UP10000063780; GB: AE004437; NID: g10580495; P1
 C: Genetics:
 A: Gene: noxC
 C: Superfamily: nitroreductase

Query Match 6.4%; Score 109.5; DB 2; Length 213;
 Best Local Similarity 20.1%; Pred. No. 0.044;
 Matches 66; Conservative 27; Mismatches 83; Indels 153; Gaps 14;

Qy 2 PDMWTTDVKSAAVOLACRAPSLHNSQPMRWIAEDHTVALFLDKDR-----VLATDHS 55
 Db 28 PDADIDDDTLAEILRDLATLAPSSYNQPMWFEVAVQ-----DDRLAEIAIAYDQH- 79
 Qy 56 GREALLCGAVLDHFRVMAAAGTTANVERPNNDPLHLASIDPSPADFTGHRLLAD 115
 Db 80 ----VDAG-----TALVAGHT-----DP-----KTADRVF----- 103
 Qy 116 ALLRRTRLPFAEPPDMDLVESQRTVTATDVRIDVIADMRPELAASKLTESRLY 175
 Db 104 -----EM-----VDGR-----F 111
 Qy 176 DSSYHAEFLFWMTGAFETSEGIPIHSSLVSAESDRTVFGDRFPVANTDRRPERGHRSKY 235
 Db 112 DADTGAEI-----KSQTVASYESDQA--GRUYAL----- 138
 Qy 236 LVSTYDNERASLIRCGEMLSAVLLDATMAGLATCTTLTELHASRDVLAALIGPATP 295
 Db 139 -----RVASL-----AAQNLISAHARGLATPMSGP-----DEFAAEFAGLPADT 180
 Qy 296 QALVAVGLAPMEBPPATPRPIDEVFH 324
 Db 181 VPVVLIVGPSGGEPRRLPRRSVDEVLH 209

RESULT 10

C87319
 C:Species: Caulobacter crescentus
 C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
 C:Accession: C87319

R: Nieman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.; B.; Laub, M.T.; Deboy, R.T.; Dodson, R.J.; Dukkin, A.S.; Gwin, M.L.; Haft, D.H.; Kolon, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
 Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
 A: Title: Complete Genome Sequence of Caulobacter crescentus.
 A: Reference number: A87249; MUID: 21173698; PMID: 11259647

A: Accession: C87319
 A: Status: preliminary
 A: Molecule type: DNA
 A: Residues: 1-409 <STO>
 A: Cross-references: UNIPROT: Q9AAN2; UNIPARC: UP100000C70B6; GB: AE005673; NID: g13421759; P1
 C: Genetics:
 A: Gene: CC0565

Query Match 6.3%; Score 109; DB 2; Length 409;
 Best Local Similarity 23.2%; Pred. No. 0.12;
 Matches 78; Conservative 34; Mismatches 128; Indels 96; Gaps 14;

Qy 17 IACRAPSILHNSQPMRWIAEDHTVALFLDKDRVLYAT-----DHSGREALIGCG---AVL 67
 Db 35 LSAVAYELKKGGDGR-----AHALSRSVGVATVRLQTSRGKALVGRPLPASAE 86

QY 68 DHRVMAAAGTTA-NVERPNDPDLASIDFSPADFTVTEGRLRADAILLRTDRLP 126
 DB 87 DGLVLTLPALSGAANLDDFDHP-----GPRL-----IVLPKWTAMP 123
 QY 127 FAEPDMDLVESQRLTATTTATDVRIDVADMDRPELAASKUTE--SLRLVSSVHAELF 184
 DB 124 DPNKKGW-----ALATRLVNDASALVLPADLRKCLTLRAPPEVGPRLTRPD--GOLF 176
 QY 185 -----WMTGAFETSEGISPHSLVSAESDRVTFGRDPFVAVNTDRRPEFG 229
 DB 177 GLQSVRAPRSLSEKGMTPVVKDQGG-----RMVLAMHTESQTVLAEPDLNTALTKLE 232
 QY 230 HDRSKVTLVSTYDNRASLRCGEMLSAVLLDATMAGLATCTLTHITELHASRDLVAALI 289
 DB 233 GARTVALLDLIHAED-----TPVFPDATHGFA-----PARSLRLILL 271
 QY 290 GQAPTPQALVRV-----GLAEMSEPPPTPRRI 319
 DB 272 EPLVGATLVLAVALAGLQAGVRFGLKERRAV 307

RESULT 11

A87128
 conserved hypothetical protein ML1751 [imported] - Mycobacterium leprae
 C:Species: Mycobacterium leprae
 C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
 C:Accession: A87128
 R:Coile, S.T.; Eigmeier, K.; Pakhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Ho
 R.; Davies, R.M.; Devlin, K.; Duthey, S.; Feltsell, T.; Fraser, A.; Hamlin, N.; Holroyd,
 eam, M.A.; Rutherford, K.M.
 Nature 409, 1007-1011, 2001
 A:Authors: Ruter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; Sq
 A:Title: Massive gene decay in the leprosy bacillus.
 A:Reference number: A86909; MUID:21128732; PMID:11234002
 A:Accession: A87128
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-721 <STO>
 A:Cross-references: UNIPROT:O9CBP5; UNIPARC:UP100000C6DE0; GB:AL450380; NID:gl3093490; F
 C:Genetics:
 A:Gene: ML1751

Query Match 6.3%; Score 108; DB 2; Length 721;
 Best Local Similarity 25.1%; Pred. No. 0.32;
 Matches 89; Conservative 34; Mismatches 103; Indels 128; Gaps 18;
 QY 10 VIKSAVQLACRAPSLSHNSQPR-----WIAEDHTVALFLDKRVLATDHSRGA 59
 DB 379 VIATA--AIRAPSGNLQPMHTIDAGPDALVIRIAPHTDSL-----DVGFSGS 424
 QY 60 LLGGAVLDHFRVMAAAGT-----TANVERFP-----NPN-----DPLHL-- 95
 DB 425 AVANGALLFNARVMAAAGVLTGPVSLAENVGAPRLVLSLGDGKNPDLAAVYEMLVRE 484
 QY 96 -----ASIDSPADFTVTEGRL-----RAD-----AILLRTRDLRFAEPDWM 133
 DB 485 TRNNHKGPTWVLTETIDVLAHAAVEGARLHLTRRDDIVRAATILAAADSRVYTRPLH 544
 QY 134 DLVESQLATTYTA--DTRIVDIADMR-PELAASKU--TESLRLYSSVHAELFWWTG 188
 DB 545 EEMISELFWPGSFPDPT-GIDVRSLKMDWHQLAAPDVLRRTDVMELAQ-----NWVG 596
 QY 189 AFETSEGISPHSLVSAESDRVTFGRDPFVAVNTDRRPEFGDR-----SKVLVSTYDNE 244
 DB 597 A-----ALGQD-----THDRVASSALAVISVPGHA 622
 QY 245 RASLRCCEMLSAVLLDATMAGLATCTLT-----HITELHASRDLVAALIGQ 291
 DB 623 LADVARGSAVVAIVAAQRYGFAVQPISPVFLYAHSDVDELHLSAQFADELGQ 676

conserved hypothetical protein Atu5016 [imported] - Agrobacterium tumefaciens (strain C5
 C:Species: Agrobacterium tumefaciens
 C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
 C:Accession: AG3161
 R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monke, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L
 erage, G.; Gillet, M.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell
 Science 294, 2317-2323, 2001
 A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
 ser, E.W.
 A:Title: The genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
 A:Reference number: AB2577; MUID:21608550; PMID:11743193
 A:Accession: AG3161
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-382 <KOR>
 A:Cross-references: UNIPROT:O8UKT3; UNIPARC:UP100000D1489; GB:AB008687; PIDN:AA145709.1;
 A:Experimental source: strain C58 (Dupont)
 A:Genetics:
 A:Gene: Atu5016
 C:Supfamily: Rhizobium plasmid pNGR234a yacF protein

Query Match 6.1%; Score 104.5; DB 2; Length 382;
 Best Local Similarity 21.5%; Pred. No. 0.27;
 Matches 73; Conservative 36; Mismatches 114; Indels 117; Gaps 15;
 QY 69 HFRVMAAAGTTANVERPNDPDLASIDFSPA--DFVTEGRLRADAILLRTDRLP 126
 DB 48 HFTACSLWVDGL-----VHLEDVLHDTATDITPTHELTIARDVLTERRRIA 98
 QY 127 FAEPDMDLVESQRLT-----TVTADTVRIDVIAD----- 157
 DB 99 -AGPPDWTLSADGLRTIRQSELSASTVGAEETAAALRVVTDPEGEDEVDSENLPG 157
 QY 158 -----MRPELAASKUTE-SLELYSSY--HAEFLWTAFTSE 194
 DB 158 VDLEAIDAVLARSEAIASATPRGAGSRAATEKDPVYDLDWDEEARLDEMGRVLRQAQ 217
 QY 195 GIP-----HSSLVSAESDRVTFPG-RDPPV 219
 DB 218 ELPAVLDQAIYALDAMNELSVQHPWIGRLAAILHAGVTSQAHIAIYLGLKTFV- 276
 QY 220 ANTDRRPEFGHRSKVLSTYDNERASLLRCGE---MLSAVLLDATMAGLATCT-LTH 274
 DB 277 ---DRRRH--RDRSRLLAIHGFIAAEIGLKHDRLLIARTWVERKLDGRTSSKLP 331
 QY 275 ITELHASRDLVAALIGQA-----TPQALVRVGLAPEMER 309
 DB 332 LVLELVMAKPIVSA--GMVAKTLDVTPQARRRIVLEGLRE 369

RESULT 13

T34908
 redv protein - Streptomyces coelicolor
 C:Species: Streptomyces coelicolor
 C:Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 09-Jul-2004
 C:Accession: T34908
 R:Olliver, K.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
 submitted to the EMBL Data Library, January 1998
 A:Reference number: 221558
 A:Accession: T34908
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-395 <OLI>
 A:Cross-references: UNIPROT:O54145; UNIPARC:UP100000DAB97; EMBL:AL021409; PIDN:CAA16173.1
 A:Experimental source: strain A3(2)
 C:Genetics:
 A:Gene: redv; SCQEDB:SC3F7.02c

Query Match 6.0%; Score 102.5; DB 2; Length 395;
 Best Local Similarity 21.8%; Pred. No. 0.42;
 Matches 86; Conservative 43; Mismatches 166; Indels 99; Gaps 16;

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: March 23, 2006, 04:56:43 ; Search time 84.6574 Seconds
(without alignments)
2758.529 Million cell updates/sec

Title: US-10-617-038-25

Perfect score: 1718

Sequence: 1 MPDMMVTVDVKSAYQLACR.....PATPRRPIDVEVHVRKXDRH 331

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1718	100.0	331	053476_MYCTU	053476 mycobacteri
2	1718	100.0	333	07D712_MYCTU	07D712 mycobacteri
3	1708	99.4	331	07VET0_MYCBO	07VET0 mycobacteri
4	915	53.3	348	07D623_MYCTU	07D623 mycobacteri
5	912	53.1	344	005800_MYCTU	005800 mycobacteri
6	904	52.6	344	07X74_MYCBO	07X74 mycobacteri
7	847.5	49.3	333	07X260_MYCBO	07X260 mycobacteri
8	629	36.6	333	05YV27_NOCFA	05YV27 nocardi
9	477.5	27.8	335	05YV27_NOCFA	05YV27 nocardi
10	434.5	25.3	324	073U03_MYCBA	073U03 mycobacteri
11	417.5	24.3	332	073257_MYCTU	073257 mycobacteri
12	417	24.3	332	07X70_MYCBO	07X70 mycobacteri
13	417	24.3	344	07D627_MYCTU	07D627 mycobacteri
14	334.5	19.5	343	063H05_MYCBO	063H05 mycobacteri
15	334.5	19.5	343	063H05_MYCBO	063H05 mycobacteri
16	295	17.2	330	09R103_STRCO	09R103 streptomyc
17	266.5	15.5	375	04NVM0_9DELT	04NVM0 anaeromyxob
18	252	14.7	309	09R133_STRCO	09R133 streptomyc
19	171	10.0	335	09R144_STRCO	09R144 streptomyc
20	169	9.8	382	04IXJ3_AZOVI	04IXJ3 azotobacter
21	163	9.5	382	05LTM6_SILPO	05LTM6 silicobacter
22	156.5	9.1	393	08TND4_METAC	08TND4 methanobact
23	155.5	9.1	330	07MVP5_PORGI	07MVP5 porphyromon
24	147	8.6	352	05W242_SERMA	05W242 serratia ma
25	143	8.3	715	Y1355_MYCTU	Y1355 mycobacteri
26	143	8.3	715	07U074_MYCBO	07U074 mycobacteri
27	140.5	8.2	989	07ZM05_DESVA	07ZM05 desulfovibr
28	125	7.3	366	089RM5_BRALH	089RM5 bradyrhizob
29	124	7.2	357	05NKK0_AZOSE	05NKK0 azaricus sp
30	121	7.0	372	07D7A9_MYCBO	07D7A9 mycobacteri
31	121	7.0	372	07TYS8_MYCBO	07TYS8 mycobacteri

32	121	7.0	372	2	P95233_MYCTU	P95233 mycobacteri
33	120	7.0	449	1	COFE_MYCBA	073UJ3 mycobacteri
34	118.5	6.9	713	2	05YX2_NOCFA	05YX2 nocardi
35	113	6.6	589	2	Q4NL3_3M1CC	Q4NL3 arthrobacte
36	109.5	6.4	213	2	09H024_HALSA	09H024 halobacteri
37	109	6.3	409	2	Q9AAN2_CAUCR	Q9AAN2 caulobacteri
38	108	6.3	721	2	Q9CBP5_MYCLB	Q9CBP5 mycobacteri
39	106.5	6.2	413	2	Q9X5U2_STRLA	Q9X5U2 streptomyc
40	104.5	6.1	382	2	Q8UKT3_AGRIS	Q8UKT3 agrobacteri
41	104	6.1	380	2	Q4INB9_9BURK	Q4INB9 burkholderi
42	103	6.0	528	2	Q6G1C8_BAROU	Q6G1C8 bartonella
43	102.5	6.0	395	2	Q54145_STRCO	Q54145 streptomyc
44	102	5.9	441	2	Q8ZCH4_STRAM	Q8ZCH4 streptomyc
45	101.5	5.9	306	2	Q4T081_98PHN	Q4T081 erythrobact

ALIGNMENTS

RESULT 1						
ID	053476_MYCTU	PRELIMINARY;	PRT;	331	AA.	
AC	053476					
DT	01-JUN-1998 (TREMBlrel. 06, Created)					
DT	01-MAR-2002 (TREMBlrel. 20, Last sequence update)					
DT	01-JUN-2003 (TREMBlrel. 24, Last annotation update)					
DE	Hypothetical protein acg.					
GN	Name=acg; OrderedLocNames=Rv2032;					
OS	Mycobacterium tuberculosis.					
OC	Bacteria; Actinobacteria; Actinobacteriales;					
OC	Corynebacterineae; Mycobacteriaceae; Mycobacterium;					
OX	NCB1_TaxID=1773;					
RN	[1]					
RP	NUCLEOTIDE SEQUENCE.					
RC	STRAIN=H37RV;					
RC	MEDLINE=96295987; PubMed=9634230; DOI=10.1038/31159;					
RA	Cole S.T., Brosch R., Parkhill J., Garnier T., Churche C.M.,					
RA	Harris D.E., Gordon S.V., Eigmeier K., Gae S., Barry C.E. III,					
RA	Tekala F., Badcock K., Basham D., Brown D., Chillingworth T.,					
RA	Connor R., Davies R.M., Devlin K., Krogan A., McLean S., Hailin N.,					
RA	Holroyd S., Hornsby T., Jørgen K., Krogan A., McLean S., Hailin N.,					
RA	Murphy L.D., Oliver S., Osborne J., Quail M.A., Rajandream M.A.,					
RA	Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squares R.,					
RA	Suleston J.E., Taylor K., Whitehead S., Barrell B.G.;					
RT	"Deciphering the biology of Mycobacterium tuberculosis from the					
RT	complete genome sequence."					
RL	Nature 393:537-544 (1998).					
DR	EMBL; BX842578; CAA17246.1; -; Genomic_DNA.					
DR	PIR; G70942; G70942.					
DR	TubercuList; Rv2032; -.					
SK	Complete proteome; Hypothetical protein.					
SK	SEQUENCE 331 Aa; 36559 MM; 8FF2A825CE90C6B CRC64;					
Query Match						
Best Local Similarity 100.0%; Pred. No. 1e-133;						
Matches 331; Conservative 0; Mismatches 0; Indels 0; Gaps 0;						
QY	1	MPDMMVTVDVKSAYQLACRAPSILNSOPMWMIEDHTVALFLDKDRVLVATDHSGRAL	60			
DB	1	MPDMMVTVDVKSAYQLACRAPSILNSOPMWMIEDHTVALFLDKDRVLVATDHSGRAL	60			
QY	61	LGCGAVLDHPRVMAAAGTTANVERFPNPDLHLASIDFSPADFVTGHRRLADAILR	120			
DB	61	LGCGAVLDHPRVMAAAGTTANVERFPNPDLHLASIDFSPADFVTGHRRLADAILR	120			
QY	121	RTDRLPFAEPDMDLVESQRLTATYADTVRIDVADMDRPELAASKLTESLRVDSYH	180			
DB	121	RTDRLPFAEPDMDLVESQRLTATYADTVRIDVADMDRPELAASKLTESLRVDSYH	180			
QY	181	AELFWMTGAFETSGIPHSLSVSAESDRVTFGSDFPVANTDRRPERGHRSKVLVIST	240			
DB	181	AELFWMTGAFETSGIPHSLSVSAESDRVTFGSDFPVANTDRRPERGHRSKVLVIST	240			

QY 241 YDNERASLLRCGEMTSAVLDDATMAGLATCTLTHITELHASRDVLAALIGOPATPOLVLR 300
DB 241 YDNERASLLRCGEMTSAVLDDATMAGLATCTLTHITELHASRDVLAALIGOPATPOLVLR 300
QY 301 VGLAPEMEBPPATPRRPIDEVFHVRAKDR 331
DB 301 VGLAPEMEBPPATPRRPIDEVFHVRAKDR 331

RESULT 2

Q7D7L2_MYCTU
ID Q7D7L2_MYCTU PRELIMINARY; PRT; 333 AA.

AC Q7D7L2;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DE 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=MT2091;
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium tuberculosis complex.
OX NCBI_TaxId=1773;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CDC 1551 / Oshkosh;
RX MEDLINE=22206494; PubMed=12218036;
RX DOI=10.1128/JB.184.19.5479-5490.2002;
RA Fleischman R.D., Alland D., Eisen J.A., Carpenter L., White O.,
Peterson J.D., DeBoy R.T., Dodson R.W., Gwinn M.L., Haft D.H.,
Hickey E.K., Kolony J.F., Nelson W.C., Umayam L.A., Brimbleton M.D.,
Salzberg S.L., Delcher A., Uterback T.R., Weidman J.F., Khouri H.M.,
Gill J., Mikula A., Bishai W., Jacobs W.R. Jr., Venter J.C.,
Fraser C.M.;
RA "Whole-genome comparison of Mycobacterium tuberculosis clinical and
laboratory strains."
RL J. Bacteriol. 184:5479-5490 (2002).
DR EMBL; AE000516; AAK46370.1; -; Genomic_DNA.
KW TIGR; MT2091; -;
KW Hypothetical protein.
SQ SEQUENCE 333 AA; 36803 MW; D873842445EAD5ED CRC64;

Query Match 100.0%; Score 1718; DB 2; Length 333;
Best Local Similarity 100.0%; Pred. No. 1e-133;
Matches 331; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPDMMTTDVTKSAVQLACRAPSLHNSQPMRWIAEDHTVALFLDKDRVLYATDHSGRAL 60
DB 3 MPDMMTTDVTKSAVQLACRAPSLHNSQPMRWIAEDHTVALFLDKDRVLYATDHSGRAL 62
QY 61 LGCGAVLDHFRVMAAAGTTANVERFPNNDPLHLASIDFSPADPVTGHRRLADAILLR 120
DB 61 LGCGAVLDHFRVMAAAGTTANVERFPNNDPLHLASIDFSPADPVTGHRRLADAILLR 122
QY 121 RTDRLPFAEPDMDLVESQRLRTVYADTVRIDVADMRPELAAASKLTSRLYDSSYH 180
DB 121 RTDRLPFAEPDMDLVESQRLRTVYADTVRIDVADMRPELAAASKLTSRLYDSSYH 182
QY 123 RTDRLPFAEPDMDLVESQRLRTVYADTVRIDVADMRPELAAASKLTSRLYDSSYH 182
DB 123 RTDRLPFAEPDMDLVESQRLRTVYADTVRIDVADMRPELAAASKLTSRLYDSSYH 182
QY 181 AELFWMTGAFETSGIPHSLSVSAESDRVTFGRDFPVVANTDRRPFEGHRSKVLVLT 240
DB 181 AELFWMTGAFETSGIPHSLSVSAESDRVTFGRDFPVVANTDRRPFEGHRSKVLVLT 242
QY 183 AELFWMTGAFETSGIPHSLSVSAESDRVTFGRDFPVVANTDRRPFEGHRSKVLVLT 242
DB 183 AELFWMTGAFETSGIPHSLSVSAESDRVTFGRDFPVVANTDRRPFEGHRSKVLVLT 242
QY 241 YDNERASLLRCGEMTSAVLDDATMAGLATCTLTHITELHASRDVLAALIGOPATPOLVLR 300
DB 241 YDNERASLLRCGEMTSAVLDDATMAGLATCTLTHITELHASRDVLAALIGOPATPOLVLR 302
QY 301 VGLAPEMEBPPATPRRPIDEVFHVRAKDR 331
DB 301 VGLAPEMEBPPATPRRPIDEVFHVRAKDR 333

RESULT 3

Q7VET0_MYCBO
ID Q7VET0_MYCBO PRELIMINARY; PRT; 331 AA.
AC Q7VET0;
DT 01-OCT-2003 (TREMBlrel. 25, Created)
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Hypothetical protein Mb2056.
GN OrderedLocusNames=Mb2056;
OS Mycobacterium bovis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium tuberculosis complex.
OX NCBI_TaxId=1765;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AF2122/97;
RX MEDLINE=23709107; PubMed=12788972; DOI=10.1073/pnas.1130426100;
RA Garnier T., Eiglmeler K., Camus J.-C., Medina N., Mansoor H.,
Pryor W., Duchoy S., Grondin S., Lacroix C., Monempe C., Simon S.,
Harris B., Aikin R., Doggett J., Mayes R., Keating J., Wheeler P.R.,
Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RA "The complete genome sequence of Mycobacterium bovis."
RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882 (2003).
DR EMBL; BX248341; CAD96911.1; -; Genomic_DNA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 331 AA; 36575 MW; 8FF2A825CEBBAEB CRC64;

Query Match 99.4%; Score 1708; DB 2; Length 331;
Best Local Similarity 99.7%; Pred. No. 6.9e-133;
Matches 330; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MPDMMTTDVTKSAVQLACRAPSLHNSQPMRWIAEDHTVALFLDKDRVLYATDHSGRAL 60
DB 1 MPDMMTTDVTKSAVQLACRAPSLHNSQPMRWIAEDHTVALFLDKDRVLYATDHSGRAL 60
QY 61 LGCGAVLDHFRVMAAAGTTANVERFPNNDPLHLASIDFSPADPVTGHRRLADAILLR 120
DB 61 LGCGAVLDHFRVMAAAGTTANVERFPNNDPLHLASIDFSPADPVTGHRRLADAILLR 120
QY 121 RTDRLPFAEPDMDLVESQRLRTVYADTVRIDVADMRPELAAASKLTSRLYDSSYH 180
DB 121 RTDRLPFAEPDMDLVESQRLRTVYADTVRIDVADMRPELAAASKLTSRLYDSSYH 180
QY 123 RTDRLPFAEPDMDLVESQRLRTVYADTVRIDVADMRPELAAASKLTSRLYDSSYH 180
DB 123 RTDRLPFAEPDMDLVESQRLRTVYADTVRIDVADMRPELAAASKLTSRLYDSSYH 180
QY 181 AELFWMTGAFETSGIPHSLSVSAESDRVTFGRDFPVVANTDRRPFEGHRSKVLVLT 240
DB 181 AELFWMTGAFETSGIPHSLSVSAESDRVTFGRDFPVVANTDRRPFEGHRSKVLVLT 240
QY 183 AELFWMTGAFETSGIPHSLSVSAESDRVTFGRDFPVVANTDRRPFEGHRSKVLVLT 240
DB 183 AELFWMTGAFETSGIPHSLSVSAESDRVTFGRDFPVVANTDRRPFEGHRSKVLVLT 240
QY 241 YDNERASLLRCGEMTSAVLDDATMAGLATCTLTHITELHASRDVLAALIGOPATPOLVLR 300
DB 241 YDNERASLLRCGEMTSAVLDDATMAGLATCTLTHITELHASRDVLAALIGOPATPOLVLR 300
QY 301 VGLAPEMEBPPATPRRPIDEVFHVRAKDR 331
DB 301 VGLAPEMEBPPATPRRPIDEVFHVRAKDR 331

RESULT 4

Q7D629_MYCTU
ID Q7D629_MYCTU PRELIMINARY; PRT; 348 AA.

AC Q7D629;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DE 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=MT3212;
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium tuberculosis complex.
OX NCBI_TaxId=1773;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CDC 1551 / Oshkosh;

```

RX MEDLINE=22206494; PubMed=12218036;
RX DOI=10.1128/JB.184.19.5479-5490.2002;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J.D., DeBoy R.T., Dodson R.J., Gwinn M.L., Haft D.H.,
RA Hickey E.K., Kolonay J.F., Nelson W.C., Umayam L.A., Esmolaeva M.D.,
RA Salzberg S.L., Delcher A., Uterback T.R., Weidman J.F., Khouri H.M.,
RA Gill J., Mikula A., Bishai W., Jacobs W.R. Jr., Venter J.C.,
RA Fraser C.M.;
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL J. Bacteriol. 184:5479-5490(2002).
DR EMBL, AE000516; AAK47552.1; -; Genomic_DNA.
DR TIGR, MT3212; -.
KW Hypothetical protein.
SQ
SEQUENCE 348 AA; 38976 MW; AE78834535008BDA CRC64;

Query Match 53.3%; Score 915; DB 2; Length 348;
Best Local Similarity 56.7%; Pred. No. 3.3e-67;
Matches 182; Conservative 39; Mismatches 96; Indels 4; Gaps 2;

QY 10 VIKSAVOLACRAPSLHNSQPMRWIAE--DH-TVALFLDKDRVLATDHSGREALLGCGA 65
D 5 VIKRAVILACRAPSVHNSQPMRWVABSGSEHTTVHLFVNHRRTVPATDHSGRQAIIISCGA 64
QY 66 VLDHFRVMAAAGTTANVERFPNPNDLHLASIDFSPADFTVTEGRLRADAILLRTRDL 125
D 65 VLDHLRIAMTRAHMQANITTFEPQNPQDLATVEFSPIDHTVAGQRRAQAIIQRTRDL 124
QY 126 PFAEPPMDLVESQLRTTYADTVRIADVADMRPELAASAKLTESLRLYDSYHAELFW 185
D 125 PFDSPMYWHLFEPALRDVADKDVAMLDVSDQRTRLVVASQLSEVLRDDPYHAELFW 184
QY 186 WTGAFTSEGIPIHSLVSAESDRVTGDRPPVYANTDRPEFGHDSKVLVSTYDNER 245
D 185 WTSPEVLAHGVPPPTLASDARLRVLDGRDPVRSYQNRRAELADDSKVLVSTPSTDR 244
QY 246 ASILRCGEMTSAVLLDATTAGLCTTLTHITELHASRDLVAALIGPATPOALVRVGLAP 305
D 245 ADALRCGEVSLTILLECTMAGMATCTTLTHLIESDSRDIYGLTRQGEPOALIRVGIAF 304
QY 306 EMEPPPATPRRPIDVEVHVR 326
D 305 PLAAPVAPTPRRPRIDSVLQIR 325

RESULT 5
O05800_MYCTU PRELIMINARY; PRT; 344 AA.
AC O05800;
DT 01-JUL-1997 (TReMBLrel. 04, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN OrderedLocustNames=RV3127;
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium tuberculosis complex.
OX NCBI_TaxID=1773;
RN
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=H37Rv;
RC MEDLINE=98295987; PubMed=9634230; DOI=10.1038/31159;
RA Cole S.T., Broesch R., Parkhill J., Garnier T., Churcher C.M.,
RA Harris D.E., Gordon S.V., Eigmeier K., Gae S., Barry C.E. III,
RA Tekala F., Badcock K., Baaham D., Brown D., Chillingworth T.,
RA Connor R., Davies R.M., Devlin K., Fellwell T., Gentles S., Hamlin N.,
RA Holroyd S., Hornsby T., Jagels K., Krogh A., McLean J., Moule S.,
RA Murphy L.D., Oliver S., Osborne J., Quail M.A., Rajandream M.A.,
RA Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Stinson J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
complete genome sequence.";
```

```

RL Nature 393:537-544(1998).
DR EMBL, BX482582; CAB08361.1; -; Genomic_DNA.
DR PIR, F70922; F70922.
DR Tuberculist; RV3127; -.
KW Complete proteome; Hypothetical protein.
SQ
SEQUENCE 344 AA; 38520 MW; 9DACAA3495F03770 CRC64;

Query Match 53.1%; Score 912; DB 2; Length 344;
Best Local Similarity 56.4%; Pred. No. 5.8e-67;
Matches 181; Conservative 40; Mismatches 96; Indels 4; Gaps 2;

QY 10 VIKSAVOLACRAPSLHNSQPMRWIAE--DH-TVALFLDKDRVLATDHSGREALLGCGA 65
D 1 MLKRAVILACRAPSVHNSQPMRWVABSGSEHTTVHLFVNHRRTVPATDHSGRQAIIISCGA 60
QY 66 VLDHFRVMAAAGTTANVERFPNPNDLHLASIDFSPADFTVTEGRLRADAILLRTRDL 125
D 61 VLDHLRIAMTRAHMQANITTFEPQNPQDLATVEFSPIDHTVAGQRRAQAIIQRTRDL 120
QY 126 PFAEPPMDLVESQLRTTYADTVRIADVADMRPELAASAKLTESLRLYDSYHAELFW 185
D 121 PFDSPMYWHLFEPALRDVADKDVAMLDVSDQRTRLVVASQLSEVLRDDPYHAELFW 180
QY 186 WTGAFTSEGIPIHSLVSAESDRVTGDRPPVYANTDRPEFGHDSKVLVSTYDNER 245
D 181 WTSPEVLAHGVPPPTLASDARLRVLDGRDPVRSYQNRRAELADDSKVLVSTPSTDR 240
QY 246 ASILRCGEMTSAVLLDATTAGLCTTLTHITELHASRDLVAALIGPATPOALVRVGLAP 305
D 241 ADALRCGEVSLTILLECTMAGMATCTTLTHLIESDSRDIYGLTRQGEPOALIRVGIAF 300
QY 306 EMEPPPATPRRPIDVEVHVR 326
D 301 PLAAPVAPTPRRPRIDSVLQIR 321

RESULT 6
Q7TX74_MYCBO PRELIMINARY; PRT; 344 AA.
AC Q7TX74;
DT 01-OCT-2003 (TReMBLrel. 25, Created)
DT 01-OCT-2003 (TReMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Hypothetical protein Mb3150.
GN OrderedLocustNames=Mb3150;
OS Mycobacterium bovis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium tuberculosis complex.
OX NCBI_TaxID=1765;
RN
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AF2122/97;
RC MEDLINE=22709107; PubMed=12788972; DOI=10.1073/pnas.1130426100;
RA Garnier T., Eigmeier K., Camus J.-C., Medina N., Mansoor H.,
RA Pryor M., Duchoy S., Gordin S., Lacroix C., Monsempo C., Simon S.,
RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RT "The complete genome sequence of Mycobacterium bovis.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
DR EMBL, BX248345; CAD95242.1; -; Genomic_DNA.
KW Complete proteome; Hypothetical protein.
SQ
SEQUENCE 344 AA; 38476 MW; 9DADPB3495F24770 CRC64;

Query Match 52.6%; Score 904; DB 2; Length 344;
Best Local Similarity 56.1%; Pred. No. 2.6e-66;
Matches 180; Conservative 40; Mismatches 97; Indels 4; Gaps 2;

QY 10 VIKSAVOLACRAPSLHNSQPMRWIAE--DH-TVALFLDKDRVLATDHSGREALLGCGA 65
D 1 MLKRAVILACRAPSVHNSQPMRWVABSGSEHTTVHLFVNHRRTVPATDHSGRQAIIISCGA 60
QY 66 VLDHFRVMAAAGTTANVERFPNPNDLHLASIDFSPADFTVTEGRLRADAILLRTRDL 125
```

```

Db      61 VLDHRLIMTAAHQANITRFPQNPQDLATVEGSPIDHVAQGRNEQAIIILKRRDRL 120
Qy      126 PFAEPPMDLVESQRTTVAIDVADMPBELAAAKLTESRLTYSSYHAEFLW 185
Db      121 PFDSPMWHLFEPALROAVDQVMDVSDQRTRLVVAQLEVLRRDDPYHAELEW 180
Qy      186 WTGAFTSEGIPIHSLVSAESDRVTFGRDPVVAANTDRPEFGHDSKVLVSTYDNER 245
Db      181 WTSFPVLAHGVPPDTLASDARLKVDTGRDPVRSYQRRRAELDDRSKVLVSTSPDTR 240
Qy      246 ASILRCGEMLSAVLLDATMAGLATCTLTHITELHASRDLVALLIGQATPOLVRGGLAP 305
Db      241 ADALRCGSVSTILLECTMAGMATCTLTHLLESSDSRDYVGLTRQGEQALIRVGIALP 300
Qy      306 EMEPPATPRRPIDEVFHVH 326
Db      301 PLAAVPAPTPRRPIDSVLQIR 321

```

RESULT 7

```

Q73260 MYCPA
ID 073260 MYCPA PRELIMINARY; PRT; 333 AA.
AC 073260
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN OrderedLocustNames=MAP1743c;
OS Mycobacterium paratuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC NCBI_TaxID=1770;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=k10;
RA Li L., Bannantine J., Zhang Q., Amonsin A., Alt D., Kapur V.;
RL Submitted (SFP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AEO17233; AAS04060.1; -; Genomic_DNA.
KM Complete proteome.
SQ SEQUENCE 333 AA; 36766 MW; 8F3BA527D631E03 CRC64;

```

Query Match 49.3%; Score 847.5; DB 2; Length 333;

Best Local Similarity 52.3%; Pred. No. 1.2e-61; Matches 170; Conservative 45; Mismatches 109; Indels 1; Gaps 1;

```

Qy      8 TDVTKSAVOLACRAPSILHNSQPMRWIAEDHTVALFLDKDRLVLTDSGREALGCGAVL 67
Db      9 TEVITAAIEIACHAPSLHNSQPMRWVAGSTGVDLFVDPRTKVSADSGREAITISGAIL 68
Qy      68 DHFRVAAAAGTTANVERFPNPNDPLHLASIDFSPADFVTEGHRLLRADAILRRTRLPF 127
Db      69 DHFRVAAAAGTSSNVAGQFPNPNDPLHLASIDVPTFVVAARADLAILRRTRNLPF 128
Qy      128 AEPDMDLVESQRTTVAIDVADMPBELAAAKLTESRLTYSSYHAEFLWMT 187
Db      129 RAPGWSALBPVLADAGRDSVALDVGPDPARPLVLAARLTALRRYDDYHHELMWT 188
Qy      188 GAFETSGIRHSSLSVAESDRVTFGRDPVVAANTDRPEFGHDSKVLVSTYDNERAS 247
Db      189 SPGRKFGIPESALVSEADRDVVRFPVDPIDERSAGSYDAKILVISTPEDTRAD 248
Qy      248 LRLCGEMLSAVLLDATMAGLATCTLTHITELHASRDLVALLIGQATPOLVRGGLAPE 306
Db      249 ALKRGVLSRLILECTAGLATCPTVHTTELAGRDILQHLMDPAVPOVLIVGVEPE 308
Qy      307 MEPPATPRRPIDEVFHVAKDHR 331
Db      309 GELPPRLTPRRPLGDVLQFRPLSDR 333

```

RESULT 8

```

OSYV27 NOCPA
ID OSYV27 NOCPA PRELIMINARY; PRT; 333 AA.
AC OSYV27
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein.
GN OrderedLocustNames=nfa27670;
OS Nocardia farcinica.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Nocardia;
OC NCBI_TaxID=37329;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=IFM 10152;
RX PubMed=15466710; DOI=10.1073/pnas.0406410101;
RA Ishikawa J., Yamashita A., Mikami Y., Hoshino Y., Kurita H., Hotta K.,
RA Shiba T., Hattori M.;
RT "The complete genomic sequence of Nocardia farcinica IFM 10152."
RL Proc. Natl. Acad. Sci. U.S.A. 101:14925-14930(2004).
DR EMBL; AP006618; BAD57614.1; -; Genomic_DNA.
KM Complete proteome; Hypothetical protein.
SQ SEQUENCE 333 AA; 36173 MW; 6EEFPA9E47BD5A81 CRC64;

```

Query Match 36.6%; Score 629; DB 2; Length 333;

Best Local Similarity 44.9%; Pred. No. 1.5e-43; Matches 146; Conservative 43; Mismatches 130; Indels 6; Gaps 3;

```

Qy      1 MPDTMTTVDYKSAVOLACRAPSILHNSQPMRWIAEDHTVALFLDKDRLVLTATDSGREAL 60
Db      12 VPDQQTILAVMR-----AARAPSLHNTQPMRWFPDGRRLRLFFDDRLDADPNRGQV 67
Qy      61 LGCGAVLDHFRVAAAAGTTANVERFPNPNDPLHLASIDFSPADFVTEGHRLLRADAILLR 120
Db      68 ISCGAMLHVRTAFPAAGMHTDTKRLPAPDRPDLAALFRPWPDPDPGVRAEAIIDHR 127
Qy      121 RTRRLPFAEPPMDLVESQRTTVAIDVADMPBELAAAKLTESRLTYSSYH 180
Db      128 RTRRLPDPDPDGGGELHARKLADPHDVLVGLDTPARPLAAASEHSTALRHVMPYQ 187
Qy      181 AELFWMTGAFETSEGIPIHSLVSAESDRVTFGRDPVVAANTDRPEFGHDSKVLVST 240
Db      188 AELRWMTGHEGSHGVPAAHALASPDDEARVVPVRRPFPAGPSARRGD-QPDRGLVLST 246
Qy      241 YDNERASLRLCGEMLSAVLLDATMAGLATCTLTHITELHASRDLVALLIGQATPOLVR 300
Db      247 AGDVTSMHLAGEALSLILECTADGATCPLTHITELATTROLANLAGRAGVGVVVR 306
Qy      301 VGLAPMEPPATPRRPIDEVHV 325
Db      307 VGVAFGNEHDHP-TPRRRPDEFLTV 330

```

RESULT 9

```

OSYV8 NOCPA
ID OSYV8 NOCPA PRELIMINARY; PRT; 335 AA.
AC OSYV8
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein.
GN OrderedLocustNames=nfa28960;
OS Nocardia farcinica.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Nocardia;
OC NCBI_TaxID=37329;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=IFM 10152;
RX PubMed=15466710; DOI=10.1073/pnas.0406410101;
RA Ishikawa J., Yamashita A., Mikami Y., Hoshino Y., Kurita H., Hotta K.,
RA Shiba T., Hattori M.;
RT "The complete genomic sequence of Nocardia farcinica IFM 10152."

```

RL Proc. Natl. Acad. Sci. U.S.A. 101:14925-14930 (2004).
DR EMBL, AP006618; BAD57743.1; -; Genomic DNA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 335 AA; 36465 MW; 64236FB763B77B04 CRC64;

Query Match 27.8%; Score 477.5; DB 2; Length 335;
Best Local Similarity 38.2%; Pred. No. 5.2e-31;
Matches 125; Conservative 41; Mismatches 144; Indels 17; Gaps 8;

QY 1 MPDVTWTDVKSAYQLACRAPSLNSQPMRWIAEDHTVALFLDKRVLYATHSGREAL 60
DB 15 LPD---DSTVATATLGRAPSVHNVQPMRWIRIDRHSVHLVDPAALPRTDPDRRVV 70
QY 61 LCGGVALDHFVMAAAGTTANVERFPNPNDPLHLASIDFSPADVTEGHRLL-RADATIL 119
DB 71 ISCGAALHLYVAFNMGASAVTLPDPADENHLATIRLV--HRAICHDLAMGALAPQ 128
QY 120 RRTDR--LPPAEPDMDLVESQLRTVYADTVRIDVIADMRPELAASKLTESLR 176
DB 129 RRTDRRVAFLPLPPGY--PGLLAERAAALGAMRVVVDQPARRRLVAARAAADRHALD 185
QY 177 SSYHAELFWMTGAFETSEGISLVSAESDRVTFGDFPVVANTDRRPEFGHRSKYL 236
DB 186 PEYGFELALMSGHSSDDGVPPARNPPAPAGDELP-ARFAPRLLD--PATDLDAHL 242
QY 237 VLSTVDNERASLLRCGEMLSAVILDATWAGLATCTLTHITELHASRDIV-AALIGQAPTP 295
DB 243 VLATCDRRARLRAGEALSAVLITRNIGLATCTLEPLVEGEGRDLVRRGVALNGTMRP 302
QY 296 QALVRVGLAPEMEEPPTPRRPIDEV 322
DB 303 QAVLRVGMADAPAEPLPRTPRRPVTDV 329

RESULT 10
Q73U03 MYCPA PRELIMINARY; PRT; 324 AA.
ID Q73U03
AC Q73U03

DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DB Hypothetical protein.
GN OrderedLocusNames=MAP3273c;
OS Mycobacterium paratuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium avium complex (MAC).
OX NCBI_TaxID=1770;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=k10.
RA Li L., Bannantine J., Zhang Q., Amonsin A., Alt D., Kapur V.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE017238; AAS05821.1; -; Genomic DNA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 324 AA; 35075 MW; 6A3B229E4CBB5CD CRC64;

Query Match 25.3%; Score 434.5; DB 2; Length 324;
Best Local Similarity 36.0%; Pred. No. 1.8e-27;
Matches 118; Conservative 48; Mismatches 125; Indels 37; Gaps 9;

QY 9 DVIRSAVOLACRAPSLNSQPMRWIAEDHTVALFLDKRVLYATHSGREALGCAVLD 68
DB 9 ETVKRTALASRAFSVHNTQPMRWIRIDPAGLHLYADPAKQLPHTPDGRDLILSCGIALQ 68
QY 69 HFRVMAAAGTTANVERFPNPNDPLHLASIDFSP--ADFTVEGHRLLRADAILRTRRLP 126
DB 69 HCVTAFAAVGMSRRRLPDPDPDLHLATLEFSPTADVDVA---LAAIIRRTDRRH 125
QY 127 FAEP---DMDLVESQ-LRTTVYADTVR-----IDVIADMRPELAASKLTESLR 176
DB 126 YSCWVPVGDIALMAARAARAFAGVTLYOVVDVGLHDIYVAGSIRDHILT----- 172

QY 177 SSYHAELFWMTGAFETSEGISLVSAESDRVTFGRDPVAVANTDRRPEFGHDR 232
DB 173 QDYLAEELTWGSRFVSQVGPARNTPPDDPAKIPTRLEFGAALPMQSSA---ADDN 228
QY 233 SKVLSTYDNERASLLRCGEMLSAVILDATWAGLATCTLTHITELHASRDLVA--LIGQ 291
DB 229 AVVIALGTRHDDRRLRLAGATGAVLLTATASAGLASCPTPELETPTQTRAIVRADIFGD 288
QY 292 PATPOLVRVGLAPEMEEPPTPRRP 319
DB 289 GHFQMLRVGMARPINADPLPATPRRL 316

RESULT 11
Q73257 MYCPA PRELIMINARY; PRT; 342 AA.
ID Q73257
AC Q73257

DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DB Hypothetical protein.
GN OrderedLocusNames=MAP1746c;
OS Mycobacterium paratuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium avium complex (MAC).
OX NCBI_TaxID=1770;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=k10;
RA Li L., Bannantine J., Zhang Q., Amonsin A., Alt D., Kapur V.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE017233; AAS04063.1; -; Genomic DNA.
KW Complete proteome.
SQ SEQUENCE 342 AA; 36982 MW; 95EEA3D03FFB9DA CRC64;

Query Match 24.3%; Score 417.5; DB 2; Length 342;
Best Local Similarity 36.1%; Pred. No. 4.9e-26;
Matches 118; Conservative 45; Mismatches 125; Indels 39; Gaps 11;

QY 11 IKSAYQLACRAPSLNSQPMRWIAEDHTVALFLDKRVLYATHSGREALGCGAVLDHF 70
DB 21 IRTVIALASRAFSVHNTQPMRWIRLVGQESLHYADGGRGLPNTDPRDLILSCGALHHC 80
QY 71 RVMAAAGTTANVERFPNPNDPLHLASIDFSP-----PADFTVEGHRLLRADAILRTRRL 125
DB 81 VGFPAVGMQARVTRLPADPDHLAILELSCGTAADVDIAL-----AAIIPRRTRR 134
QY 126 FPAEP---DMDLVESQ-LRTTVYADTVRIDVIADMRPELAASKLTESLR--LYD-- 176
DB 135 HYSSMPVPGVDIALMAARAANRVTLCQVE-----DADKLHELVKQSVMDHM 181
QY 177 SSYHAELFWMTGAFETSEGISLVSAESDRVTFGRDF--FVAVANTDRRPEFGHDRS 233
DB 182 NRDLVEILAAMSGRYAAGVAPARNTPKSDGAKIP-GRIFAGPMLA-MPSGSSAAEDNS 239
QY 234 KVLVSTVDNERASLLRCGEMLSAVILDATWAGLATCTLTHITELHASRDIV-AALIGQP 292
DB 240 VLLALGTRGDRLRLRAGEATSNVLLATISMGLASCVTEPLEVADTREAIVRAEIFGGS 299
QY 293 ATPOLVRVGLAPEMEEPPTPRRP 319
DB 300 NYFQMLRLRGMPINADPLPATPRRL 326

RESULT 12
P55195 MYCTU PRELIMINARY; PRT; 332 AA.
ID P55195
AC P55195
DT 01-MAY-1997 (TReMBLrel. 03, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Hypothetical protein.

GN OrderedLocustNames=RV3131;
 OS Mycobacterium tuberculosis.
 OC Bacteria; Actinobacteriia; Actinobacteridae; Actinomycetales;
 OC Corynebacteriineae; Mycobacteriaceae; Mycobacterium;
 OC Mycobacterium tuberculosis complex.
 CX NCBI_TaxID=1773;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=H37Rv;
 RX MEDLINE=98295987; PubMed=9634230; DOI=10.1038/21159;
 RA Cole S.T., Broesch R., Parkhill J., Garnier T., Churcher C.M.,
 RA Harris D.E., Gordon S.V., Eigemeier K., Gao S., Barry C.E. III,
 RA Tekala F., Badcock K., Bauman D., Brown D., Chillingworth T.,
 RA Connor R., Davies R.M., Devlin K., Fellwell T., Gentles S., Hamlin N.,
 RA Holtroyd S., Hornby T., Jagsels K., Krogh A., McLean J., Moule S.,
 RA Murphy L.D., Oliver S., Osborne J., Quail M.A., Rajandream M.A.,
 RA Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 RA Sulston J.E., Taylor K., Whitehead S., Barrall B.G.;
 RA "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence."
 RL Nature 393:537-544(1998).
 DR EMBL; BX842582; CAB06283.1; -; Genomic_DNA.
 DR PIR; D70645; D70645.
 DR TuberculoLac; RV3131; -.
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 332 AA; 35978 MW; 8C28FF93130D131A CRC64;
 Query Match 24.3%; Score 417; DB 2; Length 332;
 Best Local Similarity 36.1%; Pred. No. 5.2e-26;
 Matches 120; Conservative 45; Mismatches 139; Indels 28; Gaps 11;
 QY 2 PDVTWTTDVIVKSAVQACRAPSLHNSQPMRWIAEDHTVALFLDKDRLVYATDHSGREALT 61
 DB 6 PD-----AETVRTVTLAVRAPSIHNTQPMRWVCPSTSELFSPRDMQLRSTDPGRRLIL 61
 QY 62 GCGAVLDHFRVMAAAGTTANVERFPNNDPLHLASIDFSP-----ADFTEGHRLRADA 116
 DB 62 SCGVALHHCVVVALSLGQAKVNRFPDPKDRCHLATIGVQPLVPDQADVAL-----AAA 115
 QY 117 ILRRTDRLEPPA--EPPDMVLVESQLRTTADTVADTVRIDVADDMRPELAASKLTESRL 174
 DB 116 IPRRTDRRAYSCMPVPGGDIALMAAPARGGVMLR-QVSLDMKRAIVAQAVLD--HV 171
 QY 175 YDSVYHAEFLFWMTGAFETSGIPHSLSVAESDRVTFGDF--PVVAN-TDRRPERGHD 231
 DB 172 TDEYRLBELTITWSGRYSGVAVGVPARNPEPSPAPIP-GRLEFAGGSLQSPEDVLP--ADD 228
 QY 232 RSKVLVSTYNBERASLIRCGEMLSAVLLDATMAGLATCTLTHTTELHASDLY-AALIG 290
 DB 229 GAATLALGTETDRLARLRAGEAASIVLLTATMAGLACCPITPELEIAKTRDAVRAEVFG 288
 QY 291 QPATPOLVRYGLAPEMEEPPATPRRIDEV 322
 DB 289 AGGYPQMLRLVGMAPINADPLPPTPRRELISOV 320
 RESULT 13
 Q7TX70_MYCBO PRELIMINARY; PRT; 332 AA.
 AC Q7TX70;
 DT 01-OCT-2003 (TREMBlrel. 25, Created)
 DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Hypothetical protein M3155.
 GN OrderedLocustNames=M3155;
 OS Mycobacterium bovis.
 OC Bacteria; Actinobacteriia; Actinobacteridae; Actinomycetales;
 OC Corynebacteriineae; Mycobacteriaceae; Mycobacterium;
 OC Mycobacterium tuberculosis complex.
 CX NCBI_TaxID=1765;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=AF2122/97;

RX MEDLINE=22709107; PubMed=12788972; DOI=10.1073/pnas.1130426100;
 RA Garnier T., Eigemeier K., Camus J.-C., Medina N., Mansoor H.,
 RA Pryor M., Duchoy S., Grondin S., Lacroix C., Monmege C., Simon S.,
 RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
 RA Parkhill J., Barrall B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
 RA "The complete genome sequence of Mycobacterium bovis."
 RT Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
 RL EMBL; BX248345; CAD95247.1; -; Genomic_DNA.
 DR Complete proteome; Hypothetical protein.
 SQ SEQUENCE 332 AA; 35978 MW; 8C28FF93130D131A CRC64;
 Query Match 24.3%; Score 417; DB 2; Length 332;
 Best Local Similarity 36.1%; Pred. No. 5.2e-26;
 Matches 120; Conservative 45; Mismatches 139; Indels 28; Gaps 11;
 QY 2 PDVTWTTDVIVKSAVQACRAPSLHNSQPMRWIAEDHTVALFLDKDRLVYATDHSGREALT 61
 DB 6 PD-----AETVRTVTLAVRAPSIHNTQPMRWVCPSTSELFSPRDMQLRSTDPGRRLIL 61
 QY 62 GCGAVLDHFRVMAAAGTTANVERFPNNDPLHLASIDFSP-----ADFTEGHRLRADA 116
 DB 62 SCGVALHHCVVVALSLGQAKVNRFPDPKDRCHLATIGVQPLVPDQADVAL-----AAA 115
 QY 117 ILRRTDRLEPPA--EPPDMVLVESQLRTTADTVADTVRIDVADDMRPELAASKLTESRL 174
 DB 116 IPRRTDRRAYSCMPVPGGDIALMAAPARGGVMLR-QVSLDMKRAIVAQAVLD--HV 171
 QY 175 YDSVYHAEFLFWMTGAFETSGIPHSLSVAESDRVTFGDF--PVVAN-TDRRPERGHD 231
 DB 172 TDEYRLBELTITWSGRYSGVAVGVPARNPEPSPAPIP-GRLEFAGGSLQSPEDVLP--ADD 228
 QY 232 RSKVLVSTYNBERASLIRCGEMLSAVLLDATMAGLATCTLTHTTELHASDLY-AALIG 290
 DB 229 GAATLALGTETDRLARLRAGEAASIVLLTATMAGLACCPITPELEIAKTRDAVRAEVFG 288
 QY 291 QPATPOLVRYGLAPEMEEPPATPRRIDEV 322
 DB 289 AGGYPQMLRLVGMAPINADPLPPTPRRELISOV 320
 RESULT 14
 Q7D627_MYCTU PRELIMINARY; PRT; 344 AA.
 ID Q7D627;
 AC Q7D627;
 DT 05-JUN-2004 (TREMBlrel. 27, Created)
 DT 05-JUN-2004 (TREMBlrel. 27, Last sequence update)
 DT 05-JUN-2004 (TREMBlrel. 27, Last annotation update)
 DE Hypothetical protein.
 GN OrderedLocustNames=MT3217;
 OS Mycobacterium tuberculosis.
 OC Bacteria; Actinobacteriia; Actinobacteridae; Actinomycetales;
 OC Corynebacteriineae; Mycobacteriaceae; Mycobacterium;
 OC Mycobacterium tuberculosis complex.
 CX NCBI_TaxID=1773;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=CDG 1551 / Oshkosh;
 RX MEDLINE=22206494; PubMed=12218036;
 RX DOI=10.1128/JB.184.19.5479-5490.2002;
 RA Fleisemann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J.D., Deboy R.T., Dodson R.J., Gaim M.U., Hatt D.H.,
 RA Hickey E.K., Kolonay J.F., Nelson W.C., Unayam L.A., Ermolaeva M.D.,
 RA Salzberg S.L., Delcher A., Uitterback T.R., Weidman J.F., Khouri H.M.,
 RA Gill J., Mikula A., Bishel W., Jacobs W.R. Jr., Venter J.C.,
 RA Fraser C.M.;
 RA "Whole-genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains."
 RL J. Bacteriol. 184:5479-5490(2002).
 DR EMBL; AE000516; AAK47555.1; -; Genomic_DNA.
 DR TIGR; MT3217; -.
 KW Hypothetical protein.
 SQ SEQUENCE 344 AA; 37048 MW; 6ACAD2A4C05DCB49 CRC64;

Query Match 24.3%; Score 417; DB 2; Length 344;
Best Local Similarity 36.1%; Pred. No. 5.5e-26;
Matches 120; Conservative 45; Mismatches 139; Indels 28; Gaps 11;

QY 2 PDVTWTVISAVOLACRAPSLHNSQPRWTAEDHTVALFLDKRVLYATDHSGREALL 61
DB 18 PD---ASTVRVLTAVRAPSIHNTQPRMRCPTSLFESRPMQSRTPDGRRLIL 73
QY 62 GCGAVLDFRVMAMAAGTTANVERPPNPNDPLHLASIFSP-----ADFTGHRRLRADA 116
DB 74 SCGVALLHHCYVALALGMOAKNRFPDPDRCHLTLTGQPLVYRQADVAL-----AAA 127
QY 117 LLRRTRDLRPPA--BPPMDLVESQLRTTVDTRIVDIADMPRELAASKLTESRL 174
DB 128 IPRRTDRRAVSCWFPVPGDDILMAARAAGVWLR-QVSLDRMKAIYAQVLD---HV 183
QY 175 YDSSYHALFVMTGAFETSEGIPIHSLVSAESDVTGGRD--PVVAN-IDRRRPFGRD 231
DB 184 TDEEYLRRLITWSGRYGAVGVPARNRPPSPSAPIP-GRLEAGDGLSQSPDVLFP--ADD 240
QY 232 RSKVLVLTYNRERASLRCGEMLSAVLLDATMAGLATCTLTHITELHASRDLY-ALIG 290
DB 241 GAIALTGTETDRLRLRAGEASIVLITRATMGLACCPTEPLEIAKTRDAVAAYVFG 300
QY 291 QPATPOLVRVGLAPEMEPPRPPATRRPIDEV 322
DB 301 AGYPOMLRVMAPINADPLPPTRRRLSQV 332

RESULT 15
Q63HY5 BURPS PRELIMINARY; PRT; 343 AA.
AC Q63HY5
DT 25-OCT-2004 (TREMBLrel. 28, Created)
DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)
DE Hypothetical protein.
GN OrderedLocustNames=BPSS2284;
OS Burkholderia pseudomallei (Pseudomonas pseudomallei).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia; pseudomallei group.
CC NCBI_TaxID=28450;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K96243;
RX PubMed=15377794; DOI=10.1073/pnas.0403302101;
RA Holden M.T.G., Tiltball R.W., Peacock S.J., Cerdano-Tarraga A.-M.,
RA Atkins T., Crossman L.C., Pitt T., Churcher C., Mungall K.L.,
RA Bentley S.D., Sebahia M., Thomson N.R., Bason N., Beacham I.R.,
RA Brooks K., Brown K.A., Brown N.F., Challis G.L., Cherevach I.,
RA Chillingworth T., Cronin A., Crosssett B., Davis P., Deshazer D.,
RA Feltwell T., Frazer A., Hance Z., Hauser H., Holroyd S., Jagers K.,
RA Keith K.E., Maddison M., Moule S., Price C., Quail M.A.,
RA Rabinowitsch E., Rutherford K., Sanders M., Simmonds M.,
RA Songstivilai S., Stevens K., Tumapa S., Vesaratchaveest M.,
RA Whitehead S., Yeats C., Barrett B.G., Oyston P.C.F., Parkhill J.,
RT "Genomic plasticity of the causative agent of melioidosis.
RL Burkholderia pseudomallei".
RU Proc. Natl. Acad. Sci. U.S.A. 101:14240-14245 (2004).
KW EMBL; BX571966; CAH39770.1; -; Genomic DNA.
DR Complete proteome; Hypothetical protein.
SQ SEQUENCE 343 AA; 37037 MW; D2391D1628436284 CRC64;

Query Match 19.5%; Score 334.5; DB 2; Length 343;
Best Local Similarity 31.9%; Pred. No. 3.7e-19;
Matches 105; Conservative 49; Mismatches 148; Indels 27; Gaps 8;

QY 9 DVTKAVOLACAPSLHNSQPRWTAEDHTVALFLDKRVLYATDHSGREALLGCGAVLD 68
DB 25 DKRLAIIHYALIAESSHNTQPRMRCPTSLFESRPMQSRTPDGRRLILSCGALL 84
QY 69 HFRVMAAAGTTANVERPPNPNDPLHLASIFSPADFTGHRRLRADAILLRRTRDLRPPA 128

DB 85 NLRVALSREGIAYVIDMFPSPPDDVIALVRLDPHGYHDSLVLPFDAILERTVTRTPYA 144
QY 129 E---PPMDLVESQLRTTVDTRIVDI-ADMRPELAASKLTESLRLYSSYHAELEF 184
DB 145 NEAIFCD---VQRLVDAAGAAEGAEIACVDAPHALDEIAELIADARLQFADPPFRRELA 201
QY 185 WMTGAFETSEGIPIHSLVSAESDVTGGRDPVAVANTDRRPERG-----H---DRSK 234
DB 202 NWVHPRRHDDMPAPFAAGVPALIDFAT-----PVVASAIRFTFDGGGLAAMHRLVDGSP 256
QY 235 VLV-LSPTVDRASLRCGEMLSAVLLDATMAGLATCTLTHITELHASRDLYAALIGOPA 293
DB 257 LIVGISASDDRDAMVAAQALRVLVAAAAGLTASYLNQPIEIDALRELRPLHLYDA 316
QY 294 TPQALVRVGLAPEMEPPRPPATRRPIDEV 322
DB 317 HPQLLRIGRGPPVAVH---APRRPLNDV 341

Search completed: March 23, 2006, 05:24:24
Job time : 86.6574 secs

THIS PAGE BLANK (USPTO)

This Page Blank (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OW protein - protein search, using sw model

Run on: March 23, 2006, 05:24:53 ; Search time 22.5753 Seconds
(without alignments)
1212.194 Million cell updates/sec

Title: US-10-617-038-25

Perfect score: 1718
Sequence: 1 MPDMMVTVDVKSAYVLACR.....PATPRRPIDVFRVAKDHR 331

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:
1: /cgn2_6/ptodata/1/1aa/5_COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/6_COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/H_COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/PCUTS_COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/RE_COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	529	30.8	104	2	US-09-072-596-244
2	529	30.8	104	2	US-09-072-967-249
3	529	30.8	104	2	US-10-193-002-244
4	529	30.8	104	2	US-10-084-843-249
5	106.5	6.2	413	2	US-09-266-965-139
6	100.5	5.8	1176	2	US-09-976-594-793
7	96.5	5.6	415	2	US-09-758-759-23
8	95	5.5	534	2	US-10-104-047-3403
9	92.5	5.4	519	2	US-09-252-991A-16969
10	90.5	5.3	664	2	US-09-252-991A-16969
11	90.5	5.3	1596	2	US-08-978-277A-4
12	90	5.2	285	2	US-09-248-766A-20009
13	89.5	5.2	371	2	US-09-902-540-15386
14	89.5	5.2	502	2	US-09-489-039A-14063
15	89	5.2	1580	1	US-08-804-227C-11
16	89	5.2	1580	1	US-08-804-198-5
17	88.5	5.2	201	2	US-09-583-110-5297
18	88.5	5.2	206	2	US-09-107-433-2696
19	88.5	5.2	696	2	US-09-107-532A-9330
20	88.5	5.2	1016	2	US-09-252-991A-25429
21	88.5	5.2	1346	1	US-08-635-121-2
22	88.5	5.2	1346	2	US-08-978-277A-2
23	87.5	5.1	732	2	US-09-252-991A-30074
24	87.5	5.1	979	2	US-09-323-872A-55
25	87.5	5.1	979	2	US-09-072-433-9
26	87	5.1	638	2	US-09-477-962-105
27	86.5	5.0	497	1	US-08-075-193-4

28	86.5	5.0	497	1	US-08-564-090A-4	Sequence 4, Appli
29	86.5	5.0	497	4	PCT-US94-06698-4	Sequence 4, Appli
30	86.5	5.0	522	2	US-09-949-016-9740	Sequence 9740, Ap
31	86.5	5.0	522	2	US-09-949-016-9741	Sequence 9741, Ap
32	86.5	5.0	3546	2	US-09-679-279-13	Sequence 13, Appl
33	86	5.0	484	2	US-09-605-703B-2520	Sequence 2520, Ap
34	85.5	5.0	646	2	US-09-328-352-6017	Sequence 6017, Ap
35	85.5	5.0	942	2	US-09-252-991A-30681	Sequence 30681, A
36	85.5	5.0	3413	2	US-10-042-665A-8	Sequence 8, Appli
37	85	4.9	3782	2	US-09-105-537-4	Sequence 4, Appli
38	84.5	4.9	869	2	US-09-489-039A-11429	Sequence 11429, A
39	83.5	4.9	479	2	US-09-328-352-5396	Sequence 5396, Ap
40	83	4.8	297	1	US-08-738-944-51	Sequence 51, Appl
41	83	4.8	297	2	US-09-263-352-41	Sequence 41, Appl
42	83	4.8	691	2	US-09-758-759-159	Sequence 159, App
43	83	4.8	737	2	US-09-602-787A-182	Sequence 182, App
44	82.5	4.8	370	2	US-08-369-822C-2	Sequence 2, Appli
45	82.5	4.8	370	2	US-08-582-776C-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-09-072-596-244
Sequence 244, Application US/09072596
Patent No. 6458366
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Skeiky, Yael A.W.
APPLICANT: Dillon, David C.
APPLICANT: Campos-Neto, Antonia
APPLICANT: Houghton, Raymond
APPLICANT: Vedrick, Thomas S.
APPLICANT: Twardzik, Daniel R.
APPLICANT: Lodes, Michael J.
APPLICANT: Hendrickson, Ronald C.
NUMBER OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF
TUBERCULOSIS
NUMBER OF SEQUENCES: 350
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/072,596
FILING DATE: 05-MAY-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Makl, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.117C9
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 244:
SEQUENCE CHARACTERISTICS:
LENGTH: 104 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-072-596-244
Query Match 30.8%; Score 529; DB 2; Length 104;
Best local similarity 100.0%; Pred. No. 86-51;

Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 74 MAAAGTTANVERFPNPNDPLHLASIDFSPADFTVTEGHRRLRADAILLRRTRDLPFAEPDDW 133
Db 1 MAAAGTTANVERFPNPNDPLHLASIDFSPADFTVTEGHRRLRADAILLRRTRDLPFAEPDDW 60

Qy 134 DLVESQLRTTVTADTVRIDVADMRPELAASKLTESLRLYDS 177
Db 61 DLVESQLRTTVTADTVRIDVADMRPELAASKLTESLRLYDS 104

RESULT 2
US-09-072-967-249

; Sequence 249, Application US/09072967
; Patent No. 6592877

; GENERAL INFORMATION:

; APPLICANT: Reed, Steven G.

; APPLICANT: Skelky, Yasir A.W.

; APPLICANT: Dillon, Davin C.

; APPLICANT: Campos-Neto, Antonio

; APPLICANT: Houghton, Raymond

; APPLICANT: Vedvick, Thomas S.

; APPLICANT: Twardzik, Daniel R.

; APPLICANT: Lodes, Michael J.

; APPLICANT: Hendrickson, Ronald C.

; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY

; NUMBER OF SEQUENCES: 355

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SEED and BERRY LLP

; STREET: 6300 Columbia Center, 701 Fifth Avenue

; CITY: Seattle

; STATE: Washington

; COUNTRY: USA

; ZIP: 98104-7092

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent in Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/072,967

; FILING DATE: 05-MAY-1998

; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

; NAME: Makl, David J.

; REGISTRATION NUMBER: 31,392

; REFERENCE/DOCKET NUMBER: 210121.411C9

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (206) 622-4900

; TELEFAX: (206) 682-6031

; INFORMATION FOR SEQ ID NO: 249:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 104 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; US-09-072-967-249

Query Match 30.8%; Score 529; DB 2; Length 104;
Best Local Similarity 100.0%; Pred. No. 8e-51;
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 74 MAAAGTTANVERFPNPNDPLHLASIDFSPADFTVTEGHRRLRADAILLRRTRDLPFAEPDDW 133
Db 1 MAAAGTTANVERFPNPNDPLHLASIDFSPADFTVTEGHRRLRADAILLRRTRDLPFAEPDDW 60

Qy 134 DLVESQLRTTVTADTVRIDVADMRPELAASKLTESLRLYDS 177
Db 61 DLVESQLRTTVTADTVRIDVADMRPELAASKLTESLRLYDS 104

RESULT 3
US-10-193-002-244

; Sequence 244, Application US/10193002
; Patent No. 6949246

; GENERAL INFORMATION:

; APPLICANT: Reed, Steven G.

; APPLICANT: Skelky, Yasir A.W.

; APPLICANT: Dillon, Davin C.

; APPLICANT: Campos-Neto, Antonio

; APPLICANT: Houghton, Raymond

; APPLICANT: Vedvick, Thomas S.

; APPLICANT: Twardzik, Daniel R.

; APPLICANT: Lodes, Michael J.

; APPLICANT: Hendrickson, Ronald C.

; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF

; TUBERCULOSIS

; NUMBER OF SEQUENCES: 350

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SEED and BERRY LLP

; STREET: 6300 Columbia Center, 701 Fifth Avenue

; CITY: Seattle

; STATE: Washington

; COUNTRY: USA

; ZIP: 98104-7092

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent in Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/10/193,002

; FILING DATE: 10-Jul-2002

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/09/072,596

; FILING DATE: 05-MAY-1998

; ATTORNEY/AGENT INFORMATION:

; NAME: Makl, David J.

; REGISTRATION NUMBER: 31,392

; REFERENCE/DOCKET NUMBER: 210121.417C9

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (206) 622-4900

; TELEFAX: (206) 682-6031

; INFORMATION FOR SEQ ID NO: 244:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 104 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; SEQUENCE DESCRIPTION: SEQ ID NO: 244:

; US-10-193-002-244

Query Match 30.8%; Score 529; DB 2; Length 104;
Best Local Similarity 100.0%; Pred. No. 8e-51;
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 74 MAAAGTTANVERFPNPNDPLHLASIDFSPADFTVTEGHRRLRADAILLRRTRDLPFAEPDDW 133
Db 1 MAAAGTTANVERFPNPNDPLHLASIDFSPADFTVTEGHRRLRADAILLRRTRDLPFAEPDDW 60

Qy 134 DLVESQLRTTVTADTVRIDVADMRPELAASKLTESLRLYDS 177
Db 61 DLVESQLRTTVTADTVRIDVADMRPELAASKLTESLRLYDS 104

RESULT 4
US-10-084-843-249

; Sequence 249, Application US/10084843
; Patent No. 6962710

; GENERAL INFORMATION:

; APPLICANT: Reed, Steven G.

; APPLICANT: Skelky, Yasir A.W.

; Dillon, Davin C.
 ; Campos-Melo, Antonio
 ; Houghton, Raymond
 ; Vedrick, Thomas S.
 ; Twardzik, Daniel R.
 ; Lodes, Michael J.
 ; Hendrickson, Ronald C.
 ; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY
 ; AND DIAGNOSIS OF TUBERCULOSIS
 ; NUMBER OF SEQUENCES: 355
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: SEED and BERRY LLP
 ; STREET: 6300 Columbia Center, 701 Fifth Avenue
 ; CITY: Seattle
 ; STATE: Washington
 ; COUNTRY: USA
 ; ZIP: 98104-7092
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/10/084,843
 ; FILING DATE: 25-Feb-2002
 ; CLASSIFICATION: <unknown>
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/072,967
 ; FILING DATE: 05-MAY-1998
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Makl, David J.
 ; REGISTRATION NUMBER: 31,392
 ; REFERENCE/DOCKET NUMBER: 210121.411C9
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (206) 622-4900
 ; TELEFAX: (206) 682-6031
 ; INFORMATION FOR SEQ ID NO: 249:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 104 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 249:
 ; US-10-084-843-249
 ;
 ; Query Match 30.8%; Score 529; DB 2; Length 104;
 ; Best Local Similarity 100.0%; Pred. NO. 8e-51;
 ; Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 ;
 ; QY 74 MAAAGTANVERFPNDPLHLASIDPSPADVTEGHRRLADAILLRTDRLPFAEPDWM 133
 ; DB 1 MAAAGTANVERFPNDPLHLASIDPSPADVTEGHRRLADAILLRTDRLPFAEPDWM 60
 ; QY 134 DLVESQLRTVTADTVRIDVADMRPEIAAASKLTESIRLYDS 177
 ; DB 61 DLVESQLRTVTADTVRIDVADMRPEIAAASKLTESIRLYDS 104
 ;
 ; RESULT 5
 ; US-09-266-965-139
 ; Sequence 139, Application US/09266965
 ; Patent No. 6495348
 ; GENERAL INFORMATION:
 ; APPLICANT: Sherman, D
 ; APPLICANT: Mao, Y
 ; APPLICANT: Varoglu, M
 ; APPLICANT: He, M
 ; APPLICANT: Sheldon, P
 ; TITLE OF INVENTION: Mitomycin biosynthetic gene cluster
 ; FILE REFERENCE: 600.456US1
 ; CURRENT APPLICATION NUMBER: US/09/266,965
 ; CURRENT FILING DATE: 1999-03-12

; EARLIER APPLICATION NUMBER: US 08/624,447
 ; EARLIER FILING DATE: 1996-08-19
 ; EARLIER APPLICATION NUMBER: PCT/US94/11279
 ; EARLIER FILING DATE: 1994-10-06
 ; EARLIER APPLICATION NUMBER: US 08/133,963
 ; EARLIER FILING DATE: 1993-10-07
 ; NUMBER OF SEQ ID NOS: 145
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 139
 ; LENGTH: 413
 ; TYPE: PRT
 ; ORGANISM: Streptomyces lavendulae
 ; US-09-266-965-139
 ;
 ; Query Match 6.2%; Score 106.5; DB 2; Length 413;
 ; Best Local Similarity 24.1%; Pred. No. 0.0073;
 ; Matches 73; Conservative 32; Mismatches 111; Indels 87; Gaps 14;
 ;
 ; QY 2 PDITWTTDVIKSAVQLACRAPSLHNSQPMRWIAEDHTVALFLDKR-----VIYAT 52
 ; DB 107 PDPRHYADPRLAGVIAQRYPDRIHPLW-----NELKGFYDEDRRRMDYEGYTRLXNL 160
 ; QY 53 DH-----SGREALGCGAVLDHFRVMAAAG-----TTANVERFPNDP 92
 ; DB 161 VHAEIKRRNPRNLVGGYAVVDHDPPEADADRSLRGWGLDQSAVIRYWN----- 216
 ; QY 93 LHLASIDFSPAD---FTVEGHRRLADAILLRTDRLPFAEPDWM DLVESQLRTVTADTV 149
 ; DB 217 AHRAGADFPVVDGSSYIREGHRALPDE--FAATRK--FAVYTRK--VRSYTGLEPYWMAEW 270
 ; QY 150 RIDVIADMRP-----ELAASKLTESIRLYDS SYHAELFW-----WTG 188
 ; DB 271 YVEPPADDRPGRGDGMGEGHRTAVQATAMWRLAESGASAFVWNPQRTGACPGCLMRS 330
 ; QY 189 AFETSEG--IPHSLSVAASDRYTFGRDPP-----VANTDRPPEFGHRSKLYV 237
 ; DB 331 THLRDGGGELPMAGLLS-----RFARERPPGTAFRPVAVTCGSGDVRVALADEAAVLV 383
 ; QY 238 LST 240
 ; DB 384 VNT 386
 ;
 ; RESULT 6
 ; US-09-976-594-793
 ; Sequence 793, Application US/09976594
 ; Patent No. 6673549
 ; GENERAL INFORMATION:
 ; APPLICANT: Furness, Michael
 ; APPLICANT: Buchbinder, Jenny
 ; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
 ; FILE REFERENCE: PA-0041 US
 ; CURRENT APPLICATION NUMBER: US/09/976,594
 ; CURRENT FILING DATE: 2001-10-12
 ; PRIOR APPLICATION NUMBER: 60/240,409
 ; NUMBER OF SEQ ID NOS: 1143
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 793
 ; LENGTH: 1176
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; OTHER INFORMATION: InCyc ID No. 6673549 1806212CD1
 ; US-09-976-594-793
 ;
 ; Query Match 5.8%; Score 100.5; DB 2; Length 1176;
 ; Best Local Similarity 22.0%; Pred. No. 0.19;
 ; Matches 57; Conservative 38; Mismatches 87; Indels 77; Gaps 13;
 ;
 ; QY 86 FPND--NDPLHLA--SIDFSPADVTEGHRRLADAIL-----LRRTR 124

Db 50 PPIYPMNCRILGHTFISKCEPANGYORLKKCCLFFPGIHCMTMPKACADLKREIE 109
Qy 125 LPFAEPDWM-DLVESQRLTTVTADTVRIDVIADMRPELAA---ASK---LTESLRLYD 176
Db 110 L-YGCPFPDPBEEBEEESTSVKTEDIIKKDAKGGKSAKAAKAGSSKXQWGMISLGLSD 168
Qy 177 S---SYHAELW-----WTCAFTSEGIPIHS-----SLVSAAS 207
Db 169 EEIVKFSSEAHMWDYFPPLAIODLKRMGKLVDMRRSFITTVNPFYDSFVWQFLTLRER 228
Qy 208 DRVTFGRDPVAVNDRRPEFGHDS-----KVLVSTYNERASALLRGE- 253
Db 229 NKIFGKRYTTPSPDQGPCHDHDQGTGEGVGPQETYLKLVLEPYPSKLSGLK--GKN 286
Qy 254 --MLSAVLDDATMAGLATC 270
Db 287 IFLVAAITLRPETMGQTNC 305

RESULT 7
US-09-758-759-23
; Sequence 23, Application US/09758759
; Patent No. 6861513
; GENERAL INFORMATION:
; APPLICANT: Hosted, Thomas J.
; APPLICANT: Wang, Tim X.
; APPLICANT: Horan, Ann C.
; TITLE OF INVENTION: Evernimicin Biosynthetic Genes
; FILE REFERENCE: ID0983K US
; CURRENT APPLICATION NUMBER: US/09/758,759
; CURRENT FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: US 60/175,751
; PRIOR FILING DATE: 2000-01-12
; NUMBER OF SEQ ID NOS: 204
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 23
; LENGTH: 415
; TYPE: PRT
; ORGANISM: Micromonospora carbonacea
; FEATURE:
; OTHER INFORMATION: evdK
US-09-758-759-23

Query Match 5.6%; Score 96.5; DB 2; Length 415;
Best Local Similarity 22.8%; Pred. No. 0.095;
Matches 74; Conservative 29; Mismatches 140; Indels 81; Gaps 15;
Qy 17 LACR-ABSLSHNSQWRWIAEDHTVALPIDKRVLYATDTHSGREALLGGAVLDHFRVAMA 75
Db 51 LAAFLSPALAEQDERVVERE---LLLDVLSALAGVDRHVFVYASSGAV----- 99
Qy 76 AAGTTAANVERPNNPDLHLASIDSPADFTEGHRLRADAIIILRTDRL----- 125
Db 100 -----YTPYTWPPYHESAT--GRASAYGRA-KLRLEBELRHDRVQPVVTRLSNV 148
Qy 126 --PFAEP-PMDLVESQRLTTVTADTVRI---DVIADMRPELAAASKLTESLRLYDSS 178
Db 149 YGPFQCRPTPGYGVSHWLEATVRGEPTRLFGDPANVVRDYAVDDVTAIMEVIAQRAQGD 208
Qy 179 YHALLFWMTGAFERSEGIPIHSLSVSAASDRVTFGRDPVAVNDRRPEFGHDSKYLVL 238
Db 209 RDR---LPTVVNVGSGRP-TSLAELLQTMSTVAGRELEVI--RDVRRQFDR----- 254
Qy 239 STYDNERASALLRGEMLSAVLDDATMAGLATCTLTHTITELASRDLVAAI-----GQ 291
Db 255 -----GNWIDTTIARETLGQARISLP-----DGVQCQNEAVLITRAGGGS 296
Qy 292 PATFOALVVRGLAPEMEBRPPATP 315
Db 297 PARPSA--RLGRASGRBPPQRP 318

RESULT 8

US-10-104-047-3403
; Sequence 3403, Application US/10104047
; Patent No. 6943241
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. 6943241el full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3403
; LENGTH: 534
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-104-047-3403

Query Match 5.5%; Score 95; DB 2; Length 534;
Best Local Similarity 25.2%; Pred. No. 0.21;
Matches 41; Conservative 24; Mismatches 52; Indels 46; Gaps 6;
Qy 187 TGAFTSEGIPIHSLSVSAASDRVTFGRDPVAVNDRRPEFGHDR-SKVLVSTYNER 245
Db 97 SSAYSSSYGYPSSL-----GR-----AITDQAGFNDTLKVPGISISIQGM 140
Qy 246 ASLLRCGEMLSAV-----LDDATMAGLATCTLTHTITELHAARDVLAALIGOPATPOL 298
Db 141 TGLKIGDDLVNAIVKTVGTMLSSSGMISATNSVPPVSSAAPFTSMAAIAARPAKQPK 200
Qy 299 VR---VGLAPEMEBRPP-----ATPRPT 319
Db 201 LKPKGNVIGGSAVPPPIKNNNIGTWDEKGSVVXAPPPQPV 243

RESULT 9
US-09-252-991A-30815
; Sequence 30815, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 30815
; LENGTH: 519
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30815

Query Match 5.4%; Score 92.5; DB 2; Length 519;
Best Local Similarity 23.9%; Pred. No. 0.38;
Matches 75; Conservative 36; Mismatches 120; Indels 83; Gaps 15;
Qy 57 REALGCGAVLDHFRVMAAAGTTAN-VERFPNPD--PLHL--ASIDSPADFTTEGH 110
Db 1 RQTLBQAGVDAEAAGIPETAAGTHQRYVAFRDBDAOHFPVGLPAQGVDDLAHRDLPVGH 60
Qy 111 RL-RADAILLRDRLLPFAEPDMDL-----VESQRLTTVTADTVRIDVIADMRPEL- 162
Db 61 QVARLDRLARRTEGQPOAVFPTQVCGFGESELEACIALARRRRRLDIGAADOSEVL 120
Qy 163 -----AAASKLTESLRLYDSSYHAELFWMTGAFETSEGIPIHSLSVSAE 206
Db 121 GTQOABERTTHPFAGSAAGDALAGSSOAGLHQYFGEVVQVLDLADL-----VQALE 173

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using SW model

Run on: March 23, 2006, 06:14:33 ; Search time 69.6072 Seconds
(without alignments)
1986.885 Million cell updates/sec

Title: US-10-617-038-25

Perfect score: 1718
Sequence: 1 MPDMMVTVDVKSANVQLACR.....PATPRRPIDEVHVRADHR 331

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_A1_Main:

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1718	100.0	331	US-10-617-038-25	Sequence 25, Appl
2	915	53.3	344	US-10-617-038-36	Sequence 36, Appl
3	529	30.8	104	US-10-193-002-244	Sequence 244, App
4	529	30.8	104	US-10-084-843-249	Sequence 249, App
5	529	30.8	104	US-11-028-898-249	Sequence 249, App
6	529	30.8	104	US-11-082-005-244	Sequence 244, App
7	417	24.3	413	US-10-617-038-40	Sequence 40, Appl
8	106.5	6.2	413	US-09-953-348-139	Sequence 139, App
9	106.5	6.2	413	US-10-267-255-139	Sequence 139, App
10	104	6.1	380	US-10-282-122A-48196	Sequence 48196, A
11	102	5.9	441	US-10-156-761-12910	Sequence 12910, A
12	99.5	5.8	441	US-10-437-963-177042	Sequence 177042, A
13	98	5.7	3352	US-10-156-761-7961	Sequence 7961, Ap
14	97.5	5.7	1094	US-10-156-761-11435	Sequence 11435, A
15	97	5.6	1354	US-10-156-761-8425	Sequence 8425, Ap
16	96.5	5.6	415	US-09-758-759-23	Sequence 23, Appl
17	96.5	5.6	415	US-11-021-825-23	Sequence 23, Appl
18	96.5	5.6	423	US-10-107-431-31	Sequence 31, Appl
19	95	5.5	534	US-10-104-047-3403	Sequence 3403, Ap
20	95	5.5	543	US-10-094-749-2430	Sequence 2430, Ap
21	94.5	5.5	319	US-10-437-963-202037	Sequence 202037, A
22	94.5	5.5	813	US-10-156-761-10119	Sequence 10119, A
23	94.5	5.5	6238	US-10-343-710-71	Sequence 71, Appl
24	94	5.4	378	US-10-282-122A-67584	Sequence 67584, A
25	93.5	5.4	198	US-09-769-744A-30	Sequence 30, Appl
26	93.5	5.4	201	US-10-472-928-1134	Sequence 1134, Ap
27	93.5	5.4	745	US-10-732-923-19911	Sequence 19911, A

28	93.5	5.4	915	4	US-10-156-761-10050	Sequence 10050, A
29	93	5.4	431	4	US-10-087-192-1644	Sequence 1644, Ap
30	93	5.4	761	4	US-10-156-761-7913	Sequence 7913, Ap
31	93	5.4	932	4	US-10-282-122A-62242	Sequence 62242, A
32	92.5	5.4	218	4	US-10-156-761-10291	Sequence 10291, A
33	92.5	5.4	1360	5	US-10-732-923-70591	Sequence 20591, A
34	92.5	5.4	1444	5	US-10-732-923-17028	Sequence 17028, A
35	92.5	5.4	7349	4	US-10-314-657-46	Sequence 46, Appl
36	92.5	5.4	7349	5	US-10-473-193-46	Sequence 46, Appl
37	91.5	5.3	750	4	US-10-282-122A-62636	Sequence 62636, A
38	91.5	5.3	750	4	US-10-282-122A-64569	Sequence 64569, A
39	91.5	5.3	1016	4	US-10-156-761-10696	Sequence 10696, A
40	90.5	5.3	421	5	US-10-732-923-11413	Sequence 11413, A
41	90.5	5.3	457	4	US-10-282-122A-63748	Sequence 63748, A
42	90.5	5.3	841	4	US-10-282-122A-54009	Sequence 54009, A
43	90.5	5.3	1596	3	US-09-902-432-4	Sequence 4, Appl1
44	90	5.2	321	4	US-10-451-467A-428	Sequence 428, App
45	90	5.2	525	6	US-11-097-143-22077	Sequence 22077, A

ALIGNMENTS

```
RESULT 1
US-10-617-038-25
; Sequence 25, Application US/10617038
; Publication No. US20040057963A1
; GENERAL INFORMATION:
; APPLICANT: Andersen, Peter
; APPLICANT: Rosenkrands, Ida
; TITLE OF INVENTION: Therapeutic TB Vaccine
; FILE REFERENCE: SSI5AUSA
; CURRENT APPLICATION NUMBER: US/10/617, 038
; CURRENT FILING DATE: 2003-07-11
; PRIOR APPLICATION NUMBER: DK PA 2002 01098
; PRIOR FILING DATE: 2002-07-13
; PRIOR APPLICATION NUMBER: US 60/401, 725
; PRIOR FILING DATE: 2002-08-07
; NUMBER OF SEQ ID NOS: 187
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 25
; LENGTH: 331
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-617-038-25
```

```
Query Match 100.0%; Score 1718; DB 4; Length 331;
Best Local Similarity 100.0%; Pred. No. 3.4e-164;
Matches 331; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPDMMVTVDVKSANVQLACRAPSILNSQPMRWIADHTVALFLDKRVLYATDHSGRAL 60
DB 1 MPDMMVTVDVKSANVQLACRAPSILNSQPMRWIADHTVALFLDKRVLYATDHSGRAL 60
QY 1 GCGAVLDHFPVAAAAGTTANVERFPNPNDPLHLASIDFSPADFTGHRRLADATILR 120
DB 1 GCGAVLDHFPVAAAAGTTANVERFPNPNDPLHLASIDFSPADFTGHRRLADATILR 120
QY 121 RTDRLPFAEPDMDLVESQRTTVDYTRIDVIADDMRPPELLAASKLTESLRLYDSYH 180
DB 121 RTDRLPFAEPDMDLVESQRTTVDYTRIDVIADDMRPPELLAASKLTESLRLYDSYH 180
QY 181 AELPMWVGAFSTSGIPHSISVSAESDRVTFGGDFPVVANTDRRPERGHRSKVLVST 240
DB 181 AELPMWVGAFSTSGIPHSISVSAESDRVTFGGDFPVVANTDRRPERGHRSKVLVST 240
QY 241 YDNERASILRSGEMLSAVLDTATAGATCTLTITITELHASRDLVAALIGPATPOLVVR 301
DB 241 YDNERASILRSGEMLSAVLDTATAGATCTLTITITELHASRDLVAALIGPATPOLVVR 301
QY 301 VGLAPEWEEPPATPRRPIDEVHVRADHR 331
DB 301 VGLAPEWEEPPATPRRPIDEVHVRADHR 331
```


APPLICATION NUMBER: US/09/072,596
FILING DATE: 05-MAY-1998
ATTORNEY/AGENT INFORMATION:
NAME: Mark David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.417C9
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 244:
SEQUENCE CHARACTERISTICS:
LENGTH: 104 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 244:
US-11-082-005-244

Query Match 30.8%; Score 529; DB 6; Length 104;
Best Local Similarity 100.0%; Pred. No. 4,8e-45; Indels 0; Gaps 0;
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 74 MAAAGTTANVERFPNPNDPLHLASIDFSPADFTGHRRLADAILLRTDRLPAPPEPDW 133
Db 1 MAAAGTTANVERFPNPNDPLHLASIDFSPADFTGHRRLADAILLRTDRLPAPPEPDW 60

Qy 134 DLVSQRLTYTADTVRIDVADDMRPELAASKLTESLRLYDS 177
Db 61 DLVSQRLTYTADTVRIDVADDMRPELAASKLTESLRLYDS 104

RESULT 7

US-10-617-038-40
Sequence 40, Application US/10617038
Publication No. US20040057963A1
GENERAL INFORMATION:
APPLICANT: Andersen, Peter
APPLICANT: Rosenkrands, Ida
TITLE OF INVENTION: Therapeutic TB Vaccine
FILE REFERENCE: SS15AUSA
CURRENT APPLICATION NUMBER: US/10/617,038
CURRENT FILING DATE: 2003-07-11
PRIOR APPLICATION NUMBER: DK PA 2002 01098
PRIOR FILING DATE: 2002-07-13
PRIOR APPLICATION NUMBER: US 60/401,725
PRIOR FILING DATE: 2002-08-07
NUMBER OF SEQ ID NOS: 187
SOFTWARE: PatentIn version 3.2
SEQ ID NO 40
LENGTH: 332
TYPE: PRT
ORGANISM: Mycobacterium tuberculosis
US-10-617-038-40

Query Match 24.3%; Score 417; DB 4; Length 332;
Best Local Similarity 36.1%; Pred. No. 4.9e-33;
Matches 120; Conservative 45; Mismatches 139; Indels 28; Gaps 11;

Qy 2 PDVTWTTDIVIKSAVOLACRAPSLNSQPRWIAEDHTVALFLDKDRVLVATDHSGRRL 61
Db 6 PD-----AETVATVTLAVRAPIHNTQPMRWVCPTSELFSRDMQLRSTDPGRRLIL 61

Qy 62 GCGAVLDHFRVMAAAGTTANVERFPNPNDPLHLASIDFSP-----ADFTEGHRRLADA 116
Db 62 SCGVALLHCVVALASLGQAKVNRFPDPKDRCHLATIGVQPLVDPQADVAL-----AAA 115

Qy 117 ILARTDRLPRA--EPPMDLVESQRLTYTADTVRIDVADDMRPELAASKLTESLR 174
Db 116 IPRRTTRRAVSCMPVPGCDTALMAARAAGCVMLR-QVSLDLMKALVADAVD---HV 171

Qy 175 YDSSYHAELFWMTGAFETSEGIPIHSLVSAESDRVTFGRDF--PVVAN--TDRRPEFGHD 231

Db 172 TDEEYLRLELTWISGRYGVSVAGVPARNPEPSPAPIP-GRLEFAGPGLSQPSDVLFP--ADD 228
Qy 232 RSKVLVSTYDNDNRASLRGEMLSAVLADATMAAGLATCTLHTITELHASRDLY-ALLIG 230
Db 229 GAALLAAGTETDRLAALRAGEAASIVLLTATMGLACCPITSEPLETAKTRADVRAEVFG 288
Qy 291 QPATQALVRVGLAPEMEBPPAPTPRRPIDEV 322
Db 289 AGGYFQWLLRVGMALPIADPLPTPRRELSCV 320

RESULT 8

US-09-953-348-139
Sequence 139, Application US/09953348
Publication No. US20030134398A1
GENERAL INFORMATION:
APPLICANT: Sherman, David. H
APPLICANT: Mao, Yingqing
APPLICANT: Varoglu, Mustafa
APPLICANT: He, Min
APPLICANT: Sheldon, Paul
TITLE OF INVENTION: MITOMYCIN BIOSYNTHETIC GENE CLUSTER
FILE REFERENCE: 600,530US1
CURRENT APPLICATION NUMBER: US/09/953,348
CURRENT FILING DATE: 2001-09-12
PRIOR APPLICATION NUMBER: PCT/US00/06394
PRIOR FILING DATE: 2000-03-10
PRIOR APPLICATION NUMBER: 09/266965
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 153
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 139
LENGTH: 413
TYPE: PRT
ORGANISM: Streptomyces lavendulae
US-09-953-348-139

Query Match 6.2%; Score 106.5; DB 3; Length 413;
Best Local Similarity 24.1%; Pred. No. 0.14;
Matches 73; Conservative 32; Mismatches 111; Indels 87; Gaps 14;

Qy 2 PDVTWTTDIVIKSAVOLACRAPSLNSQPRWIAEDHTVALFLDKR-----VLVAT 52
Db 107 PDRRHVADPRLAGVLAQRYPDIRHFLW-----NELKGFYEDDRRMVYEGYTRLYNL 160

Qy 53 DH-----SGREALLGCGAVLDHFRVMAAAG-----TTANVERFPNPNDP 92
Db 161 VHAELKRRNPRNIVGGPYAVVDHDPRAEDADRSRELRCGPGELDORSADVIRYWN----- 216

Qy 93 LHLASIDFSPAD---FVTEGHRRLADAILLRTDRLPAPPEPMDLVESQRLTYTADTV 149
Db 217 AHKAGDFVVDSSYTRBSGRRAIPDE--FAATEK--FADVTRW--VRSVTGLPVMWAEM 270

Qy 150 RIDVIADDMRP-----ELAAASKLTESLRLYDSYHAELFW-----WTG 188
Db 271 YVERPAEDDRPGRDGMBGSHRTAVQATAMWRLAESGASAFYRNQRTKACPGCLMRS 330

Qy 189 AFETSG--IPHSLVSAESDRVTFGRDF-----VVANTDRRPFEGHRSKVLV 237
Db 331 THLRDGGGELPMAGLLS-----RPARPEPPTAFRPVAVVTGSGDRLVADAEAAVLV 383

Qy 238 LST 240
Db 384 VNT 386

RESULT 9

US-10-267-255-139
Sequence 139, Application US/10267255
Publication No. US20030124689A1
GENERAL INFORMATION:
APPLICANT: Sherman, D

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/084,843
FILING DATE: 25-FEB-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/072,967
FILING DATE: 05-MAY-1998
ATTORNEY/AGENT INFORMATION:
NAME: Makl, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.411C9
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 249:
SEQUENCE CHARACTERISTICS:
LENGTH: 104 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 249:
US-10-084-843-249
Query Match 30.8%; Score 529; DB 4; Length 104;
Best Local Similarity 100.0%; Pred. No. 4.8e-45; Indels 0; Gaps 0;
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 74 MAAAGTTANVERFPNPNDPLHLASIDFSPADFTVEGHRRLRADAILLRTDRLPFAEPDWM 133
DB 1 MAAAGTTANVERFPNPNDPLHLASIDFSPADFTVEGHRRLRADAILLRTDRLPFAEPDWM 60
DB 134 DLVESQLRTVTYADTVRIDVIADDMRPBELAASKLTESLRLYDS 177
DB 61 DLVESQLRTVTYADTVRIDVIADDMRPBELAASKLTESLRLYDS 104
RESULT 5
US-11-028-898-249
Sequence 249, Application US/11028898
Publication No. US20050136069A1
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
Skelky, Yasir A.W.
Dillon, Davin C.
Campos-Neto, Antonio
Houghton, Raymond
Vedvick, Thomas S.
Twardzik, Daniel R.
Lodes, Michael J.
Hendrickson, Ronald C.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY
AND DIAGNOSIS OF TUBERCULOSIS
NUMBER OF SEQUENCES: 355
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/11/028,898

FILING DATE: 03-Jan-2005
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/10/084,843
FILING DATE: 03-Jan-2005
APPLICATION NUMBER: US/09/072,967
FILING DATE: 05-MAY-1998
ATTORNEY/AGENT INFORMATION:
NAME: Makl, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.411C9
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 249:
SEQUENCE CHARACTERISTICS:
LENGTH: 104 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 249:
US-11-028-898-249
Query Match 30.8%; Score 529; DB 6; Length 104;
Best Local Similarity 100.0%; Pred. No. 4.8e-45; Indels 0; Gaps 0;
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 74 MAAAGTTANVERFPNPNDPLHLASIDFSPADFTVEGHRRLRADAILLRTDRLPFAEPDWM 133
DB 1 MAAAGTTANVERFPNPNDPLHLASIDFSPADFTVEGHRRLRADAILLRTDRLPFAEPDWM 60
QY 134 DLVESQLRTVTYADTVRIDVIADDMRPBELAASKLTESLRLYDS 177
DB 61 DLVESQLRTVTYADTVRIDVIADDMRPBELAASKLTESLRLYDS 104
RESULT 6
US-11-082-005-244
Sequence 244, Application US/11082005
Publication No. US20050181419A1
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
Skelky, Yasir A.W.
Dillon, Davin C.
Campos-Neto, Antonio
Houghton, Raymond
Vedvick, Thomas S.
Twardzik, Daniel R.
Lodes, Michael J.
Hendrickson, Ronald C.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF
TUBERCULOSIS
NUMBER OF SEQUENCES: 350
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/11/082,005
FILING DATE: 15-Mar-2005
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/10/193,002
FILING DATE: 10-Jul-2002

```
; APPLICANT: Mao, Y
; APPLICANT: Varoglu, M
; APPLICANT: He, M
; APPLICANT: Sheldon, P
; TITLE OF INVENTION: Mitomycin biosynthetic gene cluster
; FILE REFERENCE: 600,456US1
; CURRENT APPLICATION NUMBER: US/10/267,255
; CURRENT FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: US 09/266,965
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: US 08/624,447
; PRIOR FILING DATE: 1996-08-19
; PRIOR APPLICATION NUMBER: PCT/US94/11279
; PRIOR FILING DATE: 1994-10-06
; PRIOR APPLICATION NUMBER: US 08/133,963
; PRIOR FILING DATE: 1993-10-07
; NUMBER OF SEQ ID NOS: 145
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 139
; LENGTH: 413
; TYPE: PRT
; ORGANISM: Streptomyces lavendulae
US-10-267-255-139
```

```
Query Match          6.2%; Score 106.5; DB 4; Length 413;
Best Local Similarity 24.1%; Pred. No. 0.14;
Matches 73; Conservative 32; Mismatches 111; Indels 87; Gaps 14;
```

```
QY 2 PDMVTTVVYSAVOLACRAPSLSNSQPRWIAEDHTVALFLDKCR-----VLYAT 52
DB 107 PDPRIYADPARIAGYIAQRYPDIRHFLW-----NELKGFDEDRRMDYEGYTRLNYL 160
QY 53 DH-----SGREALIGCAVLDHFRVMAAAG-----TTANVERFPNPDP 92
DB 161 VHAELKRRNPRLVGGEPYAVVDHPADPADRSSELGPMGELDQRSADVLRWN----- 216
QY 93 LHLASIDSPAD---FVTEGHLRADAILRRTDRLPFAEPDMDLVESQURTYYTADTV 149
DB 217 AHKAGADPVVVDGSGYREGHRAIPDE--PAATER--FADYTRW--VRSVGLPVWMAEW 270
QY 150 RIDVIADMRP-----ELAAASKLTESLRVLSSTHAELFW-----WTG 188
DB 271 YVEPRAEDRPGRGDGEHRTAVQATAMKRLAESGASAFYVMPQRTGKACPGCLWRS 330
QY 189 AFERSEG--IPHSLSVSAASDRVYFGDFP-----VYANDRBPFGHDSKYLAV 237
DB 331 THLRDGGSELPMAGLLS-----RFAKEFPFGTAFRVAVYTCGSGDRVEALADEAAVLV 383
QY 238 LST 240
DB 384 VNT 386
```

```
RESULT 10
US-10-282-122A-48196
; Sequence 48196, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haasebeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zykling, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Foreyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
```

```
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 48196
; LENGTH: 380
; TYPE: PRT
; ORGANISM: Burkholderia cepacia
US-10-282-122A-48196
```

```
Query Match          6.1%; Score 104; DB 4; Length 380;
Best Local Similarity 22.2%; Pred. No. 0.21;
Matches 72; Conservative 41; Mismatches 118; Indels 94; Gaps 15;
```

```
QY 11 IKSAPVOLACRAPSLS-----HNSQPRWIAEDHTVALFLDKCRVYAYDHSREALIGCGA 65
DB 95 VTGIGSADRHATTAIVADHVRPEHIVQGHVPEPIIAQPGCVLRAGHT--EA--GPDF 150
QY 66 VLDFRVMAAAGTANVERFPNPDPDLHLASIDSPADFVTEGHLRADAILRRTDRL 125
DB 151 T-----ALAGLT-----PAVACE--IKDDGTMARLPDLI 179
QY 126 PFAEPDMDLVESQURTYYTADTVRIDVIADMRPELAASKLTESLRVLSSTHAELFW 185
DB 180 EFAR-----EHLKIGTIALDLQYRSRTESIIRLA-----ERTWQTAHGTFAVLVY- 226
QY 186 WTGAFFERSEGIIPHSLSVSAASDRVYFGDFPFPVAVANTDRPFGHDSKYLVLSTYDNER 245
DB 227 ---RDQPSGSPHIALVRGAPSPDV---DTPRVV-----HEPLSVL----- 260
QY 246 ASLRGCEMLSAVLLDATMAGLATCTLHITELHASRDV--VAALIGPAPFOALVRVGLA 304
DB 261 -DLLETGSTHSTWTLDAARDI-----AERDLGVILNCGDTKEHLIDVFKA 307
QY 305 PEMEPPPATPRRPID-EVFHVRK 328
DB 308 FDEEEKAAALKRRPVDKFTGIGNQ 332
```

```
RESULT 11
US-10-156-761-12910
; Sequence 12910, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
```


This Page Blank (uspto)

```
Db      2373 -----FLRSVSELHVSCHVDLTAPPAKRPV 2398
Qy      294 TPQALVRVGLAP-----EMEPPPPA 313
Db      2399 RRVDLPFYAFQPORWMLSEADVPFA 2423
```

RESULT 14

```
US-10-156-761-11435
; Sequence 11435, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 11435
; LENGTH: 1094
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-11435
```

```
Query Match      5.7%; Score 97.5; DB 4; Length 1094;
Best Local Similarity 23.5%; Pred. No. 4.4;
Matches 51; Conservative 35; Mismatches 64; Indels 67; Gaps 13;
```

```
Qy      126 PFAEPPDMDLV-----ESQRTTVTADTVRI--DVIAD-----DMRPEL-----AAA 165
Db      460 PAEDPADMDVAFAVALRRDPAEVEGAGADTVRYAEVLPGLPAGLERVELRLGVAR 519
Qy      166 SKLTESI-RLYDSSYHAELEFWTGAFTSEGISPHSLVSAESDRTVFGDFPVPVANTDR 224
Db      520 VPLTEAVDRL--AGLEKDPGWMRLYDSLAV-----DPDRLS--GLPVPPL--ADG 564
Qy      225 RPEBGHDSKVLVLTSTYNERA---SLRCEMLSAVLDDATMAGLATCTLTHTLHAS 281
Db      565 RTTIG--PROVLPTAGDGPQAPAPALRLG-----LKVHAPPEAA 601
Qy      282 RDLVALIGOPATPOAL-----VRVGLAPMEPEPPPA 313
Db      602 HPLEKLGALRTPRAVLTTPQVRAAVALASLDDEAPA 638
```

RESULT 15

```
US-10-156-761-8425
; Sequence 8425, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
```

```
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
```

```
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 8425
; LENGTH: 1354
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-8425
```

```
Query Match      5.6%; Score 97; DB 4; Length 1354;
Best Local Similarity 24.0%; Pred. No. 6.7;
Matches 78; Conservative 33; Mismatches 110; Indels 104; Gaps 20;
```

```
Qy      32 WIAED--HTVALFLDMDRV-----LYADHSGREALGCGAVLDHFRV----- 72
Db      330 WLEFDGVNHTAEIWLNGQKVGGLTYPFARSSHDYTKMLAGNGEOLA VKITPMFPFGSG 389
Qy      73 -----AMAAAGTTAVERFPNNDPLHLASI--DPSPA--DRV-----TEGHR 111
Db      390 DKGPAGLAFVDAGAT-----MNONSPYYLAASGMDMPAVRDVSGIMHVRRLSTGHA 444
Qy      112 LRADAILLRTRDRLPFAEPPDMDLVE-----SQRRTVTA--DTVRIID-----VI 154
Db      445 VIGDP--RVDTKLP--DLPTSRAEVTIVPVANNVSASAKRYVTAAFDVDRVRSVTV 499
Qy      155 ADDMRPELAASKLTESIRLYDSSYHAELEFWTGAFTSBSGP--HSLVSA--ESD 208
Db      500 AGGESADVTFAPSTYALRLRDP-----LWMPNGY-----GRPDLHDLTLTASIDGESD 550
Qy      209 RVT-----FGRDFPVPVANTDRRPEFG--HDRSKVLVLTSTYDNERASLRCGEMLSAVL 259
Db      551 RRTTRFGIRQFGYBYTVPL-----PFGSSDRYTQPVLLGVQKARYVRKIC-----L 597
Qy      260 LDATMAGLATCTLTHTLHASRDL 284
Db      598 TRATDKGFSMLTSLVSDSTSGTDL 622
```

```
Search completed: March 23, 2006, 06:28:13
Job time : 70.6072 secs
```


Db 6524 RHRDLASIDQVVDVLVLSCEPEADGPAPEATSSALRRVLEVRWMLGDARYTDARLWVL 6583
Qy 187 T-GAETSEGI PHSLVSAESDRVTFGRDPVAVNTDRRPERFGHRSKV----LVYSTY 241
Db 6584 TRRAVAVSTGDDVEDLAAAVRGRLRTAQO---ENPRLVVIDHDSDLEVLVPLGTG 6639
Qy 242 DNERASLRCGEMLSAVLLDA-----TWAGLATCTLTHITELH 279
Db 6640 EPEAA--IRAKVLVPRLVKAAVSEKAPAMDAGTVLLTGGTGLGLVA---RLVLTTH 6694
Qy 280 ASRDLVAALIGQATPOLVAVGLAPEME 308
Db 6695 GARDLVLASRGDTAFGA---VELATTELE 6720

RESULT 2
US-11-072-512-3403
; Sequence 3403, Application US/11072512
; Publication No. US20060023945A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHICO
; APPLICANT: YOSHITAKA, TSUTOMU
; APPLICANT: OTSUKA, MOTOKYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: 084335-0191
; CURRENT APPLICATION NUMBER: US/11/072, 512
; CURRENT FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: US 60/350,978
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: JP 2001-379298
; PRIOR FILING DATE: 2001-11-05
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3403
; LENGTH: 534
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-072-512-3403

Query Match 5.5%; Score 95; DB 7; Length 534;
Best Local Similarity 25.2%; Pred. No. 0.44;
Matches 41; Conservative 24; Mismatches 52; Indels 46; Gaps 6;
Qy 187 TGAEFTSEGI PHSLVSAESDRVTFGRDPVAVNTDRRPERFGHNR-SKVLVSTYDNR 245
Db 97 SSAVSSSYGYPPSSD-----GR-----ATTDAQAFGNDTSLKVGSISSIEGCM 140
Qy 246 ASLRCGEMLSAV-----LLDATMAGLATCTLTHITELHASRDLVAALIGQATPOL 298
Db 141 TGLKIGGLTAAVTKVTGTALSSSGMTSIATNSVPSAAPKPTSMALARKPAKQPK 200
Qy 299 VR---VGLAPEMERPP-----ATPRRPI 319
Db 201 LKPKGNVIGGSAAVPPPIKNNMIGTWDKGSVVKAPPTQPV 243

RESULT 3
US-11-060-005-4

; Sequence 4, Application US/11060005
; Publication No. US20050260693A1
; GENERAL INFORMATION:
; APPLICANT: Irwin H. Gelman
; TITLE OF INVENTION: TUMOR SUPPRESSOR GENE
; FILE REFERENCE: A30558-A-FWC-A-A 070156.0597
; CURRENT APPLICATION NUMBER: US/11/060,005
; CURRENT FILING DATE: 2005-02-17
; PRIOR APPLICATION NUMBER: 09/902,432
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 08/978,277
; PRIOR FILING DATE: 1997-11-25
; PRIOR APPLICATION NUMBER: 08/665,401
; PRIOR FILING DATE: 1996-06-18
; PRIOR APPLICATION NUMBER: 08/635,121
; PRIOR FILING DATE: 1996-04-19
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1596
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-11-060-005-4

Query Match 5.3%; Score 90.5; DB 7; Length 1596;
Best Local Similarity 22.7%; Pred. No. 5.5;
Matches 62; Conservative 45; Mismatches 123; Indels 43; Gaps 11;
Qy 65 AVLDHFRVMAAAAGTANVERFP---NPNDPL-HLASIDFSPADVTGHRRLRADIL 118
Db 875 AVIDGTR-----AVTSVEERSSPMISASVTEPLEHTAGAMPVEVTEKDI AETPV 928
Qy 119 LRTRDLRPAAPPMDLVESO-----LRTYTAIDVRIDVI-----ADMPPELAAS 166
Db 929 LTQT--LPEGKDADHDVTSVDFTSEAVTATETSEALRTEBEVTEASGAETTDMSAVS 986
Qy 167 KLTESRLYDSYHAELEFWMTGAFETSEGI PHSLVSAESDRVTFGRDPVAVNTDR-- 224
Db 987 QLTOSPTEBATVQGV-ESGVLDTEERQQTALIQAAVDKVEESQVPAQTQVGTG 1045
Qy 225 -----RPEFGHDSKVTVLSTYDNERASLRCGEMLSA---VLLDATMAGLATCTLTHIT 276
Db 1046 SKALEKVEEVEDESEVLA---SEKEDVMPKGPVQAGAEHLAQGSETGOATPESLEVP 1101
Qy 277 ELHASRDLVAALIGQATPOLVAVGLAPEMER 309
Db 1102 EVTADVDVHATC--QVIKIQQLMEQAVAPESSE 1132

RESULT 4
US-11-060-005-2
; Sequence 2, Application US/11060005
; Publication No. US20050260693A1
; GENERAL INFORMATION:
; APPLICANT: Irwin H. Gelman
; TITLE OF INVENTION: TUMOR SUPPRESSOR GENE
; FILE REFERENCE: A30558-A-FWC-A-A 070156.0597
; CURRENT APPLICATION NUMBER: US/11/060,005
; CURRENT FILING DATE: 2005-02-17
; PRIOR APPLICATION NUMBER: 09/902,432
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 08/978,277
; PRIOR FILING DATE: 1997-11-25
; PRIOR APPLICATION NUMBER: 08/665,401
; PRIOR FILING DATE: 1996-06-18
; PRIOR APPLICATION NUMBER: 08/635,121
; PRIOR FILING DATE: 1996-04-19
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1346
; TYPE: PRT
; ORGANISM: Rattus norvegicus

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: March 23, 2006, 06:17:23 ; Search time 7.94316 Seconds
(without alignments)
1192.746 Million cell updates/sec

Title: US-10-617-038-25
Perfect score: 1718
Sequence: 1 MPDMMTTDVTKSAVQLACR.....PATRRRPIDVEVHVRAQDHR 331

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 169630 seqs, 28622889 residues

Total number of hits satisfying chosen parameters: 169630

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA New:*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/PC1_NEW_PUB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	101.5	5.9	7102	US-11-143-980-48	Sequence 48, Appl
2	95	5.5	534	US-11-072-512-3403	Sequence 3403, Ap
3	90.5	5.3	1596	US-11-060-005-4	Sequence 4, Appl
4	88.5	5.2	1346	US-11-060-005-2	Sequence 2, Appl
5	84.5	4.9	1734	US-11-192-967-6	Sequence 6, Appl
6	84.5	4.9	1734	US-11-193-715-6	Sequence 6, Appl
7	84	4.9	324	US-11-096-568A-9021	Sequence 9021, Ap
8	84	4.9	358	US-11-096-568A-9020	Sequence 9020, Ap
9	84	4.9	445	US-11-096-568A-9023	Sequence 9023, A
10	81.5	4.7	1264	US-11-096-568A-30287	Sequence 30287, A
11	81.5	4.7	1301	US-11-096-568A-30286	Sequence 30286, A
12	81.5	4.7	1301	US-11-096-568A-30285	Sequence 30285, A
13	81	4.7	311	US-11-000-463-345	Sequence 345, App
14	81	4.7	519	US-10-330-773-519	Sequence 519, App
15	80	4.7	513	US-11-135-667-64	Sequence 64, Appl
16	79.5	4.6	511	US-11-087-099-10103	Sequence 10103, A
17	79	4.6	299	US-10-858-730-17	Sequence 17, Appl
18	78.5	4.6	556	US-11-087-099-3966	Sequence 3966, Ap
19	78	4.5	1680	US-10-517-939-362	Sequence 362, App
20	77.5	4.5	401	US-11-096-568A-17585	Sequence 17585, A
21	77.5	4.5	1308	US-11-124-367A-396	Sequence 396, App
22	77	4.5	310	US-11-096-568A-28488	Sequence 28488, A
23	77	4.5	473	US-11-087-099-7348	Sequence 7348, Ap
24	77	4.5	510	US-11-096-568A-17760	Sequence 17760, A
25	77	4.5	536	US-11-096-568A-17759	Sequence 17759, A

26	77	4.5	1323	US-10-517-939-312	Sequence 312, App
27	77	4.5	2362	US-11-086-568A-31778	Sequence 31778, A
28	77	4.5	2433	US-11-086-568A-31777	Sequence 31777, A
29	77	4.5	2535	US-11-086-568A-31776	Sequence 31776, A
30	76.5	4.5	303	US-11-087-099-9698	Sequence 9698, Ap
31	76.5	4.5	510	US-11-087-099-8796	Sequence 8796, Ap
32	76.5	4.5	937	US-11-017-550-66	Sequence 66, Appl
33	76.5	4.5	1075	US-11-100-640-12	Sequence 12, Appl
34	76.5	4.5	1250	US-11-117-465-62	Sequence 62, Appl
35	76	4.4	1351	US-11-072-512-3529	Sequence 3529, Ap
36	76	4.4	412	US-11-096-568A-5868	Sequence 5868, Ap
37	76	4.4	419	US-11-174-150-40	Sequence 40, Appl
38	76	4.4	430	US-11-096-568A-5867	Sequence 5867, Ap
39	76	4.4	431	US-10-525-674-26	Sequence 26, Appl
40	76	4.4	436	US-11-174-150-39	Sequence 39, Appl
41	76	4.4	464	US-11-096-568A-5866	Sequence 5866, Ap
42	76	4.4	523	US-11-087-099-8632	Sequence 8632, Ap
43	76	4.4	697	US-11-096-568A-19860	Sequence 19860, A
44	75.5	4.4	697	US-11-074-176-214	Sequence 214, App
45	75.5	4.4	765	US-11-087-099-7957	Sequence 7957, Ap

ALIGNMENTS

```
RESULT 1
US-11-143-980-48
; Sequence 48, Application US/11143980
; Publication No. US20050272133A1
; GENERAL INFORMATION:
; APPLICANT: He, Min
; APPLICANT: Hucul, John
; APPLICANT: Haltli, Bradley A.
; APPLICANT: Wagneaar, Melissa M.
; APPLICANT: Graziani, Edmund
; APPLICANT: Summers, Mia
; APPLICANT: Kulowski, Kerry
; TITLE OF INVENTION: Bioynthetic Gene Cluster for the Production of a Complex
; FILE REFERENCE: AM-101426US
; CURRENT APPLICATION NUMBER: US/11/143,980
; CURRENT FILING DATE: 2005-06-03
; PRIOR APPLICATION NUMBER: US 60/664,483
; PRIOR FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/576,895
; PRIOR FILING DATE: 2004-06-03
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 48
; LENGTH: 7102
; TYPE: PRT
; ORGANISM: Streptomyces sp.
US-11-143-980-48

Query Match 5.9%; Score 101.5; DB 7; Length 7102;
Best local Similarity 24.7%; Pred. No. 4;
Matches 96; Conservative 32; Mismatches 122; Indels 139; Gaps 23;

QY 32 WIAEDHVALFLDMDRLVYATDHSGRALLGCG---AVLDHFRVMAAGAGTNAVVEFPN 88
Db 6359 WIRQSE--VFLLD-----VALPEQVRGDAAKCGHPALLD--NALGIGLGA---FVN 6403
QY 89 PNDPLH-----ASIDFSPADFTVESHG-----LRADAILAR--- 120
Db 6404 EPGGAHLPFMSGVTLLAVGATAVRVTLSPAGPTVAIRADITGAPVLSIDLAMKPLA 6463
QY 121 -----RTD--RLPFAEPP-----DWDLY-----ESQLRITVTVA----- 146
Db 6464 EQRLLAAGSGRQDALFRLBWKELFVPTGATGAPRAQSGWGLGCHDEPRLTALTAAGVSPQ 6523
QY 147 ----DTVAIVDIADDM-----RREL--AAASKLTESLR--LYDSY--HAELFW 166
```

US-11-060-005-2

Query Match 5.2%; Score 88.5; DB 7; Length 1346;

Best Local Similarity 23.3%; Pred. No. 6; 8; Matches 64; Conservative 43; Mismatches 121; Indels 47; Gaps 12;

QY 65 AVLDHFVMAAAGTTANVERP-----NPNDPL-HLASIDFSPADPTEGRLRADAIL 118
 DB 488 AVIDOTR-----AVTSVEERSPSWISASVTEPLEHTAGEAMPVBEVTEKIIIEETPV 541
 QY 119 LRRDRLEPPAEPPMDLVESQ-----LRTVTADTVRIDVI-----ADDRPELAAAS 166
 DB 542 LTQT--LREGKDAHDMVTSEVDFTSEAVTATETSEALRTEVTEASGALETTDMVSAVS 599
 QY 167 KLTESLRLYDSSVYHAELFWMT--GAFETSEGI.PHSLSVAASDRVTEFGDRPVPVANTDR 224
 DB 600 QUTDS---PDTTEATPTPOVEGSGVLDTEEBERQOALQAVADKVESQVPAQTVQR 656
 QY 225 -----RPERGHRSKVLVSTYDNERASILRCGEMLSA---VLLDATMAGLATCTLTH 274
 DB 657 TGSKALKEVVEEEDSEVLA-----SEKEDVWPKGPVQDGAHLAGSETGQATPESLE 712
 QY 275 ITEHLASRDLYAALIGQATPQALVYVGLAPMEME 309
 DB 713 VPEVTADVDVHATC--QVTKLQOLMEQAVAPESSE 745

RESULT 5

US-11-192-967-6

Sequence 6, Application US/11192967

Publication No. US20050262587A1

GENERAL INFORMATION:

APPLICANT: Flannagan, Ronald D.

APPLICANT: Mathis, John P.

TITLE OF INVENTION: Novel Bt Toxin Receptors From

FILE REFERENCE: 35718/204664

CURRENT APPLICATION NUMBER: US/11/192,967

CURRENT FILING DATE: 2005-07-29

PRIOR APPLICATION NUMBER: US/09/715,909

PRIOR FILING DATE: 2000-11-17

PRIOR APPLICATION NUMBER: 60/166,285

PRIOR FILING DATE: 1999-11-18

NUMBER OF SEQ ID NOS: 11

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 6

LENGTH: 1734

TYPE: PRT

ORGANISM: Spodoptera frugiperda

US-11-192-967-6

Query Match 4.9%; Score 84.5; DB 7; Length 1734;

Best Local Similarity 21.5%; Pred. No. 23;

Matches 53; Conservative 30; Mismatches 97; Indels 67; Gaps 9;

QY 145 TADYRIDIVADDMRPELAASKLTESLRLYDSSVYHAELFW---WTGAFETSEGI--PHS 199
 DB 628 TPPIRLRSPSSPVEENVEGHPIITEGLTATDPTTADLHFEIDMNSYATKQNGPNT 687
 QY 200 S-----LVSAESDRVTEFGD-----FPVANTDRRPEFGH 230
 DB 688 ADYHGCVEILTVYPPDPNHRAGHGLVAREVSDGVTIDYEKEVELYLVVRVIDRNTVIGP 747
 QY 231 DRSKVLVLSY-----DNERASILRCGEM-----LSAVLLDATMAGLATC 270
 DB 748 DYDEAMLVTITIIDMNDMPIMADNTLQOTLREMADEGVIVGTLATDGLPLNRYRY 807
 QY 271 TLTHTEHLASRDLYAALIGQATPQALVYVGLAPMEMEPP-----PATPRRPIDEV 322
 DB 808 TWPPIKD--TPDDLAI--NYVTGQLTVNKGQALDADDPFRFYLYKVTASDKCSLDEF 862
 QY 323 FHVRAKD 329

DB 863 FPVCPDP 869

US-11-193-715-6

Sequence 6, Application US/11193715

Publication No. US20050260675A1

GENERAL INFORMATION:

APPLICANT: Flannagan, Ronald D.

APPLICANT: Mathis, John P.

TITLE OF INVENTION: Novel Bt Toxin Receptors From

FILE REFERENCE: 35718/204664

CURRENT APPLICATION NUMBER: US/11/193,715

CURRENT FILING DATE: 2005-07-29

PRIOR APPLICATION NUMBER: US/09/715,909

PRIOR FILING DATE: 2000-11-17

PRIOR APPLICATION NUMBER: 60/166,285

PRIOR FILING DATE: 1999-11-18

NUMBER OF SEQ ID NOS: 11

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 6

LENGTH: 1734

TYPE: PRT

ORGANISM: Spodoptera frugiperda

US-11-193-715-6

Query Match 4.9%; Score 84.5; DB 7; Length 1734;

Best Local Similarity 21.5%; Pred. No. 23;

Matches 53; Conservative 30; Mismatches 97; Indels 67; Gaps 9;

QY 145 TADYRIDIVADDMRPELAASKLTESLRLYDSSVYHAELFW---WTGAFETSEGI--PHS 199
 DB 628 TPPIRLRSPSSPVEENVEGHPIITEGLTATDPTTADLHFEIDMNSYATKQNGPNT 687
 QY 200 S-----LVSAESDRVTEFGD-----FPVANTDRRPEFGH 230
 DB 688 ADYHGCVEILTVYPPDPNHRAGHGLVAREVSDGVTIDYEKEVELYLVVRVIDRNTVIGP 747
 QY 231 DRSKVLVLSY-----DNERASILRCGEM-----LSAVLLDATMAGLATC 270
 DB 748 DYDEAMLVTITIIDMNDMPIMADNTLQOTLREMADEGVIVGTLATDGLPLNRYRY 807
 QY 271 TLTHTEHLASRDLYAALIGQATPQALVYVGLAPMEMEPP-----PATPRRPIDEV 322
 DB 808 TWPPIKD--TPDDLAI--NYVTGQLTVNKGQALDADDPFRFYLYKVTASDKCSLDEF 862
 QY 323 FHVRAKD 329
 DB 863 FPVCPDP 869

RESULT 7

US-11-096-568A-9021

Sequence 9021, Application US/11096568A

Publication No. US20060048240A1

GENERAL INFORMATION:

APPLICANT: Alexandrov, Nikolai et al.

TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides

FILE REFERENCE: 2750-1592PUS2

CURRENT APPLICATION NUMBER: US/11/096,568A

CURRENT FILING DATE: 2005-04-01

NUMBER OF SEQ ID NOS: 34471

SEQ ID NO 9021

LENGTH: 324

TYPE: PRT

ORGANISM: Triticum aestivum

FEATURE: NAME/KEY: misc_feature

LOCATION: (1)..(324)

OTHER INFORMATION: Cereals Seq. ID no. 13657649
US-11-096-568A-9021

Query Match 4.9%; Score 84; DB 7; Length 324;
Best Local Similarity 23.3%; Pred. No. 2.4;
Matches 66; Conservative 28; Mismatches 107; Indels 82; Gaps 13;

QY 18 ACRAPSLHNSQPMWIAEDHTVALFLDKDRVLYATDHSGREALIGCGAVLDHFRVAMAA 76
DB 4 AVKQPLHQQOPIYAV--PSHAASA---SSPALPSVPHSGTATPTPGAGG-----AA 49
QY 77 AGTTANVERFPNPNDPLHL-----ASIDFSPADFYTEGHRRLRADATLL 119
DB 50 SASFAGGRFPFGGLDVPNLKRRGGTSMIRVEAATASVQTLLEVAKATMMRRCCLPARDL 109
QY 120 RRTDRLPAEPDMDLVESQRTTVAIDTVRIDVADDMRPELAASKLTESTRLYDSSY 179
DB 110 RLIDPL-FVYPS--TVLGERAIYVNLQIRCVITADEV-----LLNLSLDSYVYQY 158
QY 180 HAELEFWMTGAFTSEGIPIHSSLSVAESDRTVFGDPFVVANTDRRPFEGHRSKVLVL- 238
DB 159 AAEL-----QRRLQRAEGDELPLF-----EF---RALELALB 187
QY 239 ---STYDNERASILRCGEMLSAVLIDATMAGLATCTLTHITEL 278
DB 188 AACSFLLDAQAAEL---EIEAYPLDLDELTSKISTLNLRKVRRL 226

RESULT 8
US-11-096-568A-9020
Sequence 9020, Application US/11096568A
Publication No. US20060048240A1
GENERAL INFORMATION:
APPLICANT: Alexandrov, Nickolai et al.
TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
FILE REFERENCE: 2750-1592PUS2
CURRENT APPLICATION NUMBER: US/11/096,568A
CURRENT FILING DATE: 2005-04-01
NUMBER OF SEQ ID NOS: 34471
SEQ ID NO 9020
LENGTH: 358
TYPE: PRT
ORGANISM: Triticum aestivum
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)_(358)
OTHER INFORMATION: Cereals Seq. ID no. 13657648
US-11-096-568A-9020

Query Match 4.9%; Score 84; DB 7; Length 358;
Best Local Similarity 23.3%; Pred. No. 2.8;
Matches 66; Conservative 28; Mismatches 107; Indels 82; Gaps 13;

QY 18 ACRAPSLHNSQPMWIAEDHTVALFLDKDRVLYATDHSGREALIGCGAVLDHFRVAMAA 76
DB 38 AVKQPLHQQOPIYAV--PSHAASA---SSPALPSVPHSGTATPTPGAGG-----AA 83
QY 77 AGTTANVERFPNPNDPLHL-----ASIDFSPADFYTEGHRRLRADATLL 119
DB 84 SASFAGGRFPFGGLDVPNLKRRGGTSMIRVEAATASVQTLLEVAKATMMRRCCLPARDL 143
QY 120 RRTDRLPAEPDMDLVESQRTTVAIDTVRIDVADDMRPELAASKLTESTRLYDSSY 179
DB 144 RLIDPL-FVYPS--TVLGERAIYVNLQIRCVITADEV-----LLNLSLDSYVYQY 192
QY 180 HAELEFWMTGAFTSEGIPIHSSLSVAESDRTVFGDPFVVANTDRRPFEGHRSKVLVL- 238
DB 193 AAEL-----QRRLQRAEGDELPLF-----EF---RALELALB 221
QY 239 ---STYDNERASILRCGEMLSAVLIDATMAGLATCTLTHITEL 278
DB 222 AACSFLLDAQAAEL---EIEAYPLDLDELTSKISTLNLRKVRRL 260

RESULT 9
US-11-096-568A-9023
Sequence 9023, Application US/11096568A
Publication No. US20060048240A1
GENERAL INFORMATION:
APPLICANT: Alexandrov, Nickolai et al.
TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
FILE REFERENCE: 2750-1592PUS2
CURRENT APPLICATION NUMBER: US/11/096,568A
CURRENT FILING DATE: 2005-04-01
NUMBER OF SEQ ID NOS: 34471
SEQ ID NO 9023
LENGTH: 445
TYPE: PRT
ORGANISM: Triticum aestivum
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)_(445)
OTHER INFORMATION: Cereals Seq. ID no. 16611691
US-11-096-568A-9023

Query Match 4.9%; Score 84; DB 7; Length 445;
Best Local Similarity 23.3%; Pred. No. 3.8;
Matches 66; Conservative 28; Mismatches 107; Indels 82; Gaps 13;

QY 18 ACRAPSLHNSQPMWIAEDHTVALFLDKDRVLYATDHSGREALIGCGAVLDHFRVAMAA 76
DB 4 AVKQPLHQQOPIYAV--PSHAASA---SSPALPSVPHSGTATPTPGAGG-----AA 49
QY 77 AGTTANVERFPNPNDPLHL-----ASIDFSPADFYTEGHRRLRADATLL 119
DB 50 SASFAGGRFPFGGLDVPNLKRRGGTSMIRVEAATASVQTLLEVAKATMMRRCCLPARDL 109
QY 120 RRTDRLPAEPDMDLVESQRTTVAIDTVRIDVADDMRPELAASKLTESTRLYDSSY 179
DB 110 RLIDPL-FVYPS--TVLGERAIYVNLQIRCVITADEV-----LLNLSLDSYVYQY 158
QY 180 HAELEFWMTGAFTSEGIPIHSSLSVAESDRTVFGDPFVVANTDRRPFEGHRSKVLVL- 238
DB 159 AAEL-----QRRLQRAEGDELPLF-----EF---RALELALB 187
QY 239 ---STYDNERASILRCGEMLSAVLIDATMAGLATCTLTHITEL 278
DB 188 AACSFLLDAQAAEL---EIEAYPLDLDELTSKISTLNLRKVRRL 226

RESULT 10
US-11-096-568A-30287
Sequence 30287, Application US/11096568A
Publication No. US20060048240A1
GENERAL INFORMATION:
APPLICANT: Alexandrov, Nickolai et al.
TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
FILE REFERENCE: 2750-1592PUS2
CURRENT APPLICATION NUMBER: US/11/096,568A
CURRENT FILING DATE: 2005-04-01
NUMBER OF SEQ ID NOS: 34471
SEQ ID NO 30287
LENGTH: 1264
TYPE: PRT
ORGANISM: Arabidopsis thaliana
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)_(1264)
OTHER INFORMATION: Cereals Seq. ID no. 4952354
US-11-096-568A-30287

Query Match 4.7%; Score 81.5; DB 7; Length 1264;
Best Local Similarity 23.1%; Pred. No. 2.9;

Best Local Similarity 100.0%; Pred. No. 1.2e-108; Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFLTTIGQPPAYQVLTALIGDLSKVDAKQPGDYTTTTSDEHGKRVVFFWPKDPTFV 60
 DB 1 MFLTTIGQPPAYQVLTALIGDLSKVDAKQPGDYTTTTSDEHGKRVVFFWPKDPTFV 60
 QY 61 CPTETIAAFSKUNDEFEDRDAQIIIGVSIIDSEFAHFQWRAQHNDLKTLPFPMISDIKRELSQ 120
 DB 61 CPTETIAAFSKUNDEFEDRDAQIIIGVSIIDSEFAHFQWRAQHNDLKTLPFPMISDIKRELSQ 120
 QY 121 AAGVLANADGVADRVTFIVDPNNEIQFVSATAGSVGRNVDEVLRVLDALQSDDELCAQWVRK 180
 DB 121 AAGVLANADGVADRVTFIVDPNNEIQFVSATAGSVGRNVDEVLRVLDALQSDDELCAQWVRK 180
 QY 181 GDPFLTDAGEELLKASA 195
 DB 181 GDPFLTDAGEELLKASA 195

RESULT 2

AAV43802
 ID AAV43802 standard; protein; 195 AA.

AC AAV43802;
 DT 11-FEB-2000 (first entry)

DE A alkyl hydroperoxide reductase subunit C protein.

KW Hydroperoxide reductase subunit C protein; resistance;
 KM antimicrobial reactive nitrogen intermediate; antibacterial;
 KM nitric oxide; nitrite; nitrosonium; S-nitrosothiol; nitrogen dioxide;
 KM dinitrogen trioxide; dinitrogen tetroxide; septic hypotension; stroke;
 KM infection; vaccine.

OS Mycobacterium tuberculosis.

PN W09954479-A2.

PD 28-OCT-1999.

PF 21-APR-1999; 99WO-US008704.

PR 21-APR-1998; 98US-0082573P.

PA (CORR) CORNELL RES FOUND INC.

PI Nathan CF, Xie Q, Chen L;

DR WPI; 2000-013258/01.

XX N-PSDB; AAZ30392.

PT Novel uses of alkyl hydroperoxide reductase subunit C (AHRIC).

PS Claim 40; Page 16-17; 74pp; English.

XX The present sequence represents an alkyl hydroperoxide reductase subunit
 CC C protein, which confers on Mycobacterium tuberculosis resistance against
 CC antimicrobial reactive nitrogen intermediates. The protein is used in a
 CC method for screening therapeutics for their ability to sensitize bacteria
 CC to antibacterial effects of reactive nitrogen intermediates. The method
 CC comprises preparing recombinant host cells lacking endogenous alkyl
 CC hydroperoxide reductase subunit C protein encoding genes, placing cells
 CC in growth medium containing reactive nitrogen intermediates and a test
 CC compound, and determining if the cells survive. The alkyl hydroperoxide
 CC reductase subunit C polypeptides can confer resistance to reactive
 CC nitrogen intermediates, e.g. nitric oxide, nitrite, nitrosonium, S-
 CC nitrosothiol, nitrogen dioxide, dinitrogen trioxide, and dinitrogen
 CC tetroxide. The methods can be used to identify agents that inhibit the
 CC reactive nitrogen intermediate-resistance mechanisms of pathogens. The
 CC alkyl hydroperoxide reductase subunit C polypeptides can also be used to
 CC inhibit production of reactive nitrogen intermediates in the treatment of

CC e.g. septic hypotension, stroke or infections. The alkyl hydroperoxide
 CC reductase subunit C protein or DNA can be used for vaccinating mammals
 CC against infection by M. tuberculosis. The antibodies against the protein
 CC can be used for passively immunising mammals infected with M.
 CC tuberculosis. These products can also be used for detection of M.
 CC tuberculosis

XX Sequence 195 AA;

SO Query Match 100.0%; Score 1024; DB 3; Length 195;

Best Local Similarity 100.0%; Pred. No. 1.2e-108;

Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFLTTIGQPPAYQVLTALIGDLSKVDAKQPGDYTTTTSDEHGKRVVFFWPKDPTFV 60
 DB 1 MFLTTIGQPPAYQVLTALIGDLSKVDAKQPGDYTTTTSDEHGKRVVFFWPKDPTFV 60
 QY 61 CPTETIAAFSKUNDEFEDRDAQIIIGVSIIDSEFAHFQWRAQHNDLKTLPFPMISDIKRELSQ 120
 DB 61 CPTETIAAFSKUNDEFEDRDAQIIIGVSIIDSEFAHFQWRAQHNDLKTLPFPMISDIKRELSQ 120
 QY 121 AAGVLANADGVADRVTFIVDPNNEIQFVSATAGSVGRNVDEVLRVLDALQSDDELCAQWVRK 180
 DB 121 AAGVLANADGVADRVTFIVDPNNEIQFVSATAGSVGRNVDEVLRVLDALQSDDELCAQWVRK 180
 QY 181 GDPFLTDAGEELLKASA 195
 DB 181 GDPFLTDAGEELLKASA 195

RESULT 3

AAM50749
 ID AAM50749 standard; protein; 195 AA.

AC AAM50749;

DT 18-APR-2002 (first entry)

DE Mycobacterium tuberculosis immunodominant Mtb protein APhc2.

KW Mtb; APhc2; immunogen; mycobacteria; immunisation; vaccine.

OS Mycobacterium tuberculosis.

PN W0200204018-A2.

PD 17-JAN-2002.

PF 10-JUL-2001; 2001WO-US021717.

PR 10-JUL-2000; 2000US-0217646P.

PA (COLS) UNIV COLORADO STATE RES FOUND.

PI Orme IM, Belisle JT;

DR WPI; 2002-164602/21.

XX The present sequence is that of the Mycobacterium tuberculosis (Mtb)
 CC strain H37Rv cytosolic Rv2428 gene product, designated APhc2. This is one
 CC of 31 immunodominant secreted or cytosolic Mtb proteins of strain H37Rv
 CC (see AAM50729-59) discovered through the use of 2-dimensional liquid
 CC phase electrophoresis coupled with an in vitro interferon-gamma assay and
 CC liquid chromatography-mass spectrometry. The immunogens stimulate a
 CC strong interferon-gamma response from T cells of M. tuberculosis infected
 CC mice. The invention provides vaccine compositions for boosting immunity
 CC to mycobacteria when administered in mid-life to a subject who has been

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: March 23, 2006, 04:56:29 ; Search time 51.844 Seconds
(without alignments) 1652.628 Million cell updates/sec

Title: US-10-617-038-26

Perfect score: 1024

Sequence: 1 MRLTIGQFPAYQLTALIG.....CMRKGDPTLDAGELTKASA 195

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1024	100.0	195	2	AAR98369 Mycobacte
2	1024	100.0	195	3	AAY43802 A alkyl h
3	1024	100.0	195	5	AAM50749 Mycobacte
4	1024	100.0	195	6	ABU34818 Protein e
5	1024	100.0	195	6	ABU36784 Protein e
6	1024	100.0	195	6	ADJ33705 M. tuberc
7	935	91.3	195	6	ABU33962 Protein e
8	926	90.4	195	6	AA67761 Amino aci
9	904	88.3	195	6	ABU36067 Protein e
10	901.5	88.0	193	2	AAR14303 M.avium a
11	898	87.7	195	7	ADB74335 Mycobacte
12	693	67.7	198	6	ABU25828 Protein e
13	566.5	55.3	179	9	AEB39863 L. pneumo
14	566.5	55.3	198	9	AEB36453 L. pneumo
15	530.5	51.8	188	8	ADS30355 Bacteri
16	524.5	51.2	184	9	ABW91466 M. xanthu
17	432.5	42.2	182	6	ABU21078 Protein e
18	410.5	40.1	182	6	ABU22889 Protein e
19	361.5	35.3	180	8	AD28683 Bacteri
20	353.5	34.5	193	8	ADS42506 Bacteri
21	351.5	34.3	192	8	ADS29794 Bacteri
22	351	34.3	199	9	AEA00256 Leishmani
23	350.5	34.2	200	8	ADM48264 Polypepti
24	350.5	34.2	200	8	ADN20134 Bacteri

25	350.5	34.2	283	5	AAE29736	AAE29736 Mycobacte
26	350.5	34.2	282	5	AAE29727	AAE29727 Leishmani
27	350.5	34.2	290	5	AAE29725	AAE29725 Leishmani
28	350.5	34.2	321	5	AAE29728	AAE29728 Leishmani
29	350.5	34.2	328	5	AAE29726	AAE29726 Leishmani
30	350.5	34.2	446	6	ABP55386	ABP55386 Human col
31	350.5	34.2	926	7	ADB78891	ADB78891 Leishmani
32	350.5	34.2	930	5	AAE29731	AAE29731 Mycobacte
33	350.5	34.2	930	7	ADA26364	ADA26364 Mycobacte
34	350.5	34.2	955	7	ADB78890	ADB78890 Leishmani
35	350.5	34.2	982	5	AAU71855	AAU71855 Leishmani
36	350.5	34.2	982	5	AAU71309	AAU71309 MASP1A-M1
37	350.5	34.2	982	7	ADB78858	ADB78858 Leishmani
38	350.5	34.2	1427	5	AAU71857	AAU71857 Leishmani
39	350.5	34.2	1427	5	AAU71311	AAU71311 MASP1A-M1
40	350.5	34.2	1427	7	ADB78860	ADB78860 Leishmani
41	350.5	34.2	1464	6	ABR54576	ABR54576 Protease
42	350.5	34.2	1464	7	ADB14458	ADB14458 Expressio
43	350.5	34.2	1464	7	ADG26981	ADG26981 Human pro
44	350.5	34.2	1641	5	AAU71856	AAU71856 Leishmani
45	350.5	34.2	1641	5	AAU71310	AAU71310 MASP1A-M1

ALIGNMENTS

RESULT 1	
AAE29736	AAE29736 standard; protein; 195 AA.
ID	AAE29736
XX	AAE29736
AC	AAE29736
XX	AAE29736
DT	16-OCT-2003 (revised)
DT	19-SEP-1996 (first entry)
XX	AAE29736
DE	Mycobacterial AhpC polypeptide.
XX	ahpC operon; AhpC, AhpF, alkyl hydroperoxide reductase; isoniazid;
KM	antibiotic resistance; mycobacteria; therapy; diagnosis; vaccine.
XX	XX
OS	Mycobacterium bovis; strain ATCC 35729.
XX	XX
PD	W09619578-A2.
XX	XX
PD	27-JUN-1996.
XX	XX
PF	20-DEC-1995; 95WO-NZ000133.
XX	XX
PR	20-DEC-1994; 94NZ-00270222.
XX	XX
PA	(NZPA-) NEW ZEALAND PASTORAL AGRIC RES INST LTD.
XX	XX
PI	Collins DW, Wilson T;
XX	XX
DR	WPI; 1996-309592/31.
XX	XX
PT	N-PSDB; AATJ1461.
XX	XX
PT	Polynucleotide(s) derived from the mycobacteria ahpC operon - useful in
XX	XX
PS	diagnosis and treatment of mycobacterial diseases.
XX	XX
XX	Example 2; Fig 2; 55pp; English.
CC	The ahpC operon (AATJ1461) of isoniazid (INH)-resistant Mycobacterium
CC	bovis ATCC encodes polypeptides AhpC (AAE29736) and AhpF (AAE29737),
CC	which likely combine to form an active alkyl hydroperoxide reductase that
CC	may be either a direct target for INH or acts to confer INH resistance.
CC	Isolated AhpC and/or AhpF polypeptides can be prep'd, e.g. in mammalian or
CC	yeast systems. They are useful for drug screening, in immunosays and in
CC	vaccine prodn. (Updated on 16-OCT-2003 to standardise OS field)
XX	XX
SO	Sequence 195 AA;
Query Match	100.0%; Score 1024; DB 2; Length 195;

Query Match 4.7%; Score 81; DB 7; Length 311;
Best Local Similarity 21.8%; Pred. No. 4.5;
Matches 69; Conservative 42; Mismatches 118; Indels 88; Gaps 16;

QY 13 SAVOLACRAPSLHNSQPMRWIAEDHTVALFLDKDRVLYATDHSREALLGCAVLDPHPRV 72
DB 31 TSVEICRSLDPQATWTF-WYRQ-----FPKQSLMATSNESKATYEQGVKDKFLI 83
QY 73 AMAAGTTANVERPNPND-PLHLASIDFSPAD-----FVTEGRLADALLARTDRLP 126
DB 84 NHAALTJSTLTVTSIAHPEDSFYICSAESTSDPKNEQFGPGRLTV-----TSDLK 136
QY 127 FAEPPD---MDLVESQRTTVAADTVRIDVIADMRPELAASKLITSLRLYDSSYAE 183
DB 137 NVFPEPVAVPEPSSAEISHTOKATLV---CLATGFPD-----HVEL 175
QY 184 FMTGAFETSEGI---PHSLVSAESD-RVTFGRDPFVANTDRP----- 226
DB 176 SMVNGKEVHSGVSTDPQLKEQPALNDSRYCLSSRLVSATFWQPNPNHRCQVQFYGL 235
QY 227 -----EFGHDSKVL--VLSTYDNERASILRG-----EMLSAVLIDATMAGLATCTL 272
DB 236 SENDEWQDRAKPYTOIVSABAMGRAD---CGFTSESYOQVLSATILYEILLGKAT--- 289
QY 273 THITELHASRDVLAALI 289
DB 290 -----LYAV--LVSAALV 299

RESULT 14

US-10-330-773-519
; Sequence 519, Application US/10330773
; Publication No. US20060040262A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 529452001300
; CURRENT APPLICATION NUMBER: US/10/330,773
; CURRENT FILING DATE: 2002-12-27
; NUMBER OF SEQ ID NOS: 981
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 519
; LENGTH: 519
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-330-773-519

Query Match 4.7%; Score 81; DB 6; Length 519;
Best Local Similarity 20.5%; Pred. No. 9.2;
Matches 53; Conservative 27; Mismatches 77; Indels 102; Gaps 12;

QY 9 DVKSAVQ-LACRAPSLHNSQPMRWIAEDHTVALFLDKDRVLY-----ATDHSGR 57
DB 166 DIISVIEENKMCNLP-NQOQPERWDEDAISRLGTGAEMTFYHIKQEPSSASHTGC 223
QY 58 EALLGCCGAVLDHFRVMAAAGTTANVERPNPNDPLHLASIDFSPADVTGSHLRDAI 117
DB 224 E---GCAA-----LVYGEVEYAEEDFSDVDAA--EV 251
QY 118 LARTRDLRFAEPPDMDLVESQRTTVAADTVR-IDVIAD-----MRPELAASK 167
DB 252 MLK--SKVPGE--DWVLPQCITLTVHPPIKDIAYVAQNFCACGCTPTQPKFVKLR 307
QY 168 LTESLRLY-----DSSYHAELFWMTGAFB 191
DB 308 YCEYLGKCYFCASCHSAESCIPARILTMWDRKXQVSDFSKMLDSVWHQVFKLGGH- 366
QY 192 TSEGIPIHSLVSAESDRV 210
DB 367 -----HSLYAKAKELDRV 379

RESULT 15
US-11-135-667-64
; Sequence 64, Application US/11135667
; Publication No. US20060034817A1
; GENERAL INFORMATION:
; APPLICANT: ROBERTS, JOSEPH
; APPLICANT: SETHURAMAN, NATARAJAN
; APPLICANT: MACALISTER, THOMAS
; TITLE OF INVENTION: CLONING, OVEREXPRESSION AND THERAPEUTIC USE OF
; FILE REFERENCE: 078728/0106
; CURRENT APPLICATION NUMBER: US/11/135,667
; PRIOR FILING DATE: 2005-05-24
; PRIOR APPLICATION NUMBER: US/09/833,745
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/197,770
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: Patentia Ver. 2.1
; SEQ ID NO 64
; LENGTH: 513
; TYPE: PRT
; ORGANISM: Halobacterium sp.
US-11-135-667-64

Query Match 4.7%; Score 80; DB 7; Length 513;
Best Local Similarity 22.1%; Pred. No. 11;
Matches 81; Conservative 36; Mismatches 95; Indels 154; Gaps 21;

QY 6 VTTDVIKSAVQ-LACRAPSLHNSQPMRWIAEDHTVALFLDKDRVLYATDHSGR--- 57
DB 229 LTTEVMTST--ASCAPAH--EPVRPHQOAVSAHINRLTGSGVLDHNRCDGRVOD 284
QY 58 EALLGC---GAV---LDHFRVMAAAGTTANVERPNPNDPLHLASIDFSPADVTG 109
DB 285 AYSIRCLPQVHGAVRDALDHLRAAVATELNSAT-----DNPL-----VFPSTVSG 331
QY 110 HRRLADAILL-----RRTRRL-----PFAEPPDMDLVESQRT 142
DB 332 GNTHGEVTLARLGYASALAEIAISERRTRDLNPTQEPLEPFLAP-----DSGLHS 365
QY 143 TVTADTVRIDVIADMRPELAASKLITSLRLYDSSYHAELFWMTGAFETSEGI 199
DB 386 GL-----MIPQYTAAS-LVNDLR-----SLGQPATLDN 412
QY 200 SLVSAESDRVTFGRDPFVANTDRRPERGHSKVLVSTYD-----NERASL-----LR 250
DB 413 ASVSGAQEDHVSMSAG-----AAVYFREAVEKATVVGVELL 449
QY 251 CGEMLSAVLIDATMAGLATCTLTHITELHASRDVLAALIGOPATPOA-----LVR 300
DB 450 CGAQ-GREFLDPLALGATA-----AAVYLVSEVSEPGDRALADMAAVGDDLV 499
QY 301 VGLAPE 306
DB 500 AGLVED 505

Search completed: March 23, 2006, 06:29:36
Job time: 8.94316 secs

[illegible][illegible]

CC vaccinated neonatally or in early childhood with BCG and in whom
CC protective immunity has waned. The vaccine compositions comprise 1 or
CC more of the 31 purified immunogenic proteins. A preferred protein is
CC Ag55A (see AAM50759), the secreted product of the Rv3084v gene
XX
SO Sequence 195 AA;

Query Match 100.0%; Score 1024; DB 5; Length 195;
Best Local Similarity 100.0%; Pred. No. 1.2e-108;
Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPLLTIQGPAPYQLTALIGDLSKVDAKQPGDYFTTITSDHSGKRVVFFWMPDFFV 60
DB 1 MPLLTIGQFPAPYQLTALIGDLSKVDAKQPGDYFTTITSDHSGKRVVFFWMPDFFV 60
QY 61 CPTETAAFSKLNDEFEDBDQAQILGVSISEFAHFQWRAQHNDLKTLPFPMLSDIKRELSQ 120
DB 61 CPTETAAFSKLNDEFEDBDQAQILGVSISEFAHFQWRAQHNDLKTLPFPMLSDIKRELSQ 120
QY 121 AAGVLMADGVADRVTFIVDPNNIEIQFSATAGSVGRNDEVLRVLDALQSDCLCAQNRK 180
DB 121 AAGVLMADGVADRVTFIVDPNNIEIQFSATAGSVGRNDEVLRVLDALQSDCLCAQNRK 180
QY 181 GDDPTLDAGELLKASA 195
DB 181 GDDPTLDAGELLKASA 195

RESULT 4
ABU34818
ID ABU34818 standard; protein; 195 AA.
XX
AC ABU34818;
XX
DT 19-JUN-2003 (first entry)
XX
DE Protein encoded by Prokaryotic essential gene #20345.
XX
KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
OS Mycobacterium bovis.
XX
PN WO200277183-A2.
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002WO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zykkind JM;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
DR WPI; 2003-029926/02.
DR N-PSDB; ACA38688.
XX
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 25; SEQ ID NO 62742; 1766bp; English.
XX
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense

CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

Sequence 195 AA;

Query Match 100.0%; Score 1024; DB 6; Length 195;
Best Local Similarity 100.0%; Pred. No. 1.2e-108;
Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPLLTIQGPAPYQLTALIGDLSKVDAKQPGDYFTTITSDHSGKRVVFFWMPDFFV 60
DB 1 MPLLTIGQFPAPYQLTALIGDLSKVDAKQPGDYFTTITSDHSGKRVVFFWMPDFFV 60
QY 61 CPTETAAFSKLNDEFEDBDQAQILGVSISEFAHFQWRAQHNDLKTLPFPMLSDIKRELSQ 120
DB 61 CPTETAAFSKLNDEFEDBDQAQILGVSISEFAHFQWRAQHNDLKTLPFPMLSDIKRELSQ 120
QY 121 AAGVLMADGVADRVTFIVDPNNIEIQFSATAGSVGRNDEVLRVLDALQSDCLCAQNRK 180
DB 121 AAGVLMADGVADRVTFIVDPNNIEIQFSATAGSVGRNDEVLRVLDALQSDCLCAQNRK 180
QY 181 GDDPTLDAGELLKASA 195
DB 181 GDDPTLDAGELLKASA 195

RESULT 5
ABU36784
ID ABU36784 standard; protein; 195 AA.
XX
AC ABU36784;
XX
DT 19-JUN-2003 (first entry)
XX
DE Protein encoded by Prokaryotic essential gene #22311.
XX
KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
OS Mycobacterium tuberculosis.
XX
PN WO200277183-A2.
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002WO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.

PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362899P.
XX
XX (ELIT-) ELITRA PHARM INC.
XX
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zykkind JW;
PI Mail D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX WPI; 2003-029926/02.
DR N-PSDB; ACA40654.
XX
XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
XX Claim 25; SEQ ID NO 64708; 1766pp; English.
XX
XX The invention relates to an isolated nucleic acid comprising any one of
CC the 623 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 195 AA;
Query Match 100.0%; Score 1024; DB 6; Length 195;
Best Local Similarity 100.0%; Pred. No. 1,2e-108;
Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MPLLITGDOPPAYQUTLALIGDLSKVDAKQPGDYFTTITSDEHFGKRWVFFPKDPTFV 60
DB 1 MPLLITIDQFPAYQUTLALIGDLSKVDAKQPGDYFTTITSDEHFGKRWVFFPKDPTFV 60
QY 61 CPTETIAAFSKLNDEFEDRDQILGVSTIDSEFAHFQWRQNHDKTLTLPFPMLSDIKRELSQ 120
DB 61 CPTETIAAFSKLNDEFEDRDQILGVSTIDSEFAHFQWRQNHDKTLTLPFPMLSDIKRELSQ 120
QY 121 AAGVLANADVADRTFTVDPNNEIQFVSATAGSVGRNVDEVLRVLDALQSDCLCACMWRK 180
DB 121 AAGVLANADVADRTFTVDPNNEIQFVSATAGSVGRNVDEVLRVLDALQSDCLCACMWRK 180
QY 181 GDPTLDAGELLKASA 195
DB 181 GDPTLDAGELLKASA 195
RESULT 6

AD137305
ID AD137305 standard; protein: 195 AA.
XX
XX AD137305;
AC
XX
XX 22-APR-2004 (first entry)
DT
XX
XX M. tuberculosis low oxygen induced antigen Rv2428 SEQ ID NO:26.
DE
XX
XX mycobacterial infection; vaccine; tuberculosis;
KW Mycobacterium tuberculosis; immunisation; antibacterial; gene therapy;
XX low oxygen induced antigen.
XX
XX Mycobacterium tuberculosis.
XX
XX WO2004006952-A2.
XX
XX 22-JAN-2004.
XX
XX 08-JUL-2003; 2003WO-DK000477.
XX
XX 13-JUL-2002; 2002DK-00001098.
XX
XX (STAT-) STATENS SERUM INSTR.
XX
XX Andersen P, Rosenkrands I, Stryhn A;
PI WPI; 2004-122778/12.
DR N-PSDB; AD137350.
XX
XX Use of one or more polypeptides or their fragments, which are expressed
PT during the latent stage of the mycobacterial infection, and/or nucleic
PT acids encoding the polypeptides, for a therapeutic vaccine against
PT tuberculosis.
XX
XX Claim 3; SEQ ID NO 26; 76pp; English.
XX
XX The present invention describes polypeptides or their fragments, which
CC are expressed during the latent stage of a mycobacterial infection,
CC and/or nucleic acids encoding the polypeptides, which are useful for
CC creating a therapeutic vaccine against tuberculosis. Also described: (1)
CC a therapeutic vaccine against tuberculosis comprising one or more
CC polypeptides; (2) a method for treating an animal, including a human
CC being, with tuberculosis caused by virulent mycobacteria, e.g. by
CC Mycobacterium tuberculosis, *M. africanum* or *M. bovis*; (3) a method for
CC immunising an animal, including a human being, against tuberculosis
CC caused by virulent mycobacteria; (4) a method of diagnosing tuberculosis
CC caused by virulent mycobacteria in an animal, including a human being;
CC (5) a method for diagnosing previous or ongoing infection with a virulent
CC mycobacterium; and (6) a method of diagnosing Mycobacterium tuberculosis
CC infection in a subject. The polypeptides have antibacterial activities,
CC and can be used in vaccines and in gene therapy. The polypeptides are
CC useful for the manufacture of a therapeutic vaccine for treating an
CC individual who is infected by a virulent mycobacterium, e.g. *M.*
CC tuberculosis, and who is not vaccinated with BCG against tuberculosis.
CC The present sequence represents a low oxygen induced antigen, which is
CC used in the exemplification of the present invention.
XX
SQ Sequence 195 AA;
Query Match 100.0%; Score 1024; DB 8; Length 195;
Best Local Similarity 100.0%; Pred. No. 1,2e-108;
Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MPLLITGDOPPAYQUTLALIGDLSKVDAKQPGDYFTTITSDEHFGKRWVFFPKDPTFV 60
DB 1 MPLLITIDQFPAYQUTLALIGDLSKVDAKQPGDYFTTITSDEHFGKRWVFFPKDPTFV 60
QY 61 CPTETIAAFSKLNDEFEDRDQILGVSTIDSEFAHFQWRQNHDKTLTLPFPMLSDIKRELSQ 120
DB 61 CPTETIAAFSKLNDEFEDRDQILGVSTIDSEFAHFQWRQNHDKTLTLPFPMLSDIKRELSQ 120
QY 121 AAGVLANADVADRTFTVDPNNEIQFVSATAGSVGRNVDEVLRVLDALQSDCLCACMWRK 180
DB 121 AAGVLANADVADRTFTVDPNNEIQFVSATAGSVGRNVDEVLRVLDALQSDCLCACMWRK 180

Db 121 AAGVLMADGVADRVTFIVDPNNIEIQFVSATAGSVGRNVDEVLRVLDALQSDDELCAQNRK 180
QY 181 GDPTLDAGELKASA 195
Db 181 GDPTLDAGELKASA 195

RESULT 7

ABU33962 standard; protein; 195 AA.

ABU33962;

19-JUN-2003 (first entry)

Protein encoded by Prokaryotic essential gene #19489.

Anti-sense; prokaryotic essential gene; cell proliferation; drug design.

Mycobacterium avium.

MO200277183-A2.

03-OCT-2002.

21-MAR-2002; 2002WO-US009107.

21-MAR-2001; 2001US-00815242.

06-SEP-2001; 2001US-00948993.

25-OCT-2001; 2001US-0342923P.

08-FEB-2002; 2002US-00072851.

06-MAR-2002; 2002US-0362699P.

(ELIT-) ELITRA PHARM INC.

Wang L, Zamudio C, Malone C, Hasebeck R, Ohlsen KL, Zyskind JW;
Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

WPI; 2003-029926/02.

N-PSDB; ACP37832.

New anti-sense nucleic acids, useful for identifying proteins or screening
for homologous nucleic acids required for cellular proliferation to
isolate candidate molecules for rational drug discovery programs.

Claim 25; SEQ ID NO 61886; 1766pp; English.

The invention relates to an isolated nucleic acid comprising any one of
the 6213 anti-sense sequences given in the specification where expression
of the nucleic acid inhibits proliferation of a cell. Also included are:
(1) a vector comprising a promoter operably linked to the nucleic acid
encoding a polypeptide whose expression is inhibited by the anti-sense
nucleic acid; (2) a host cell containing the vector; (3) an isolated
polypeptide or its fragment whose expression is inhibited by the
anti-sense nucleic acid; (4) an antibody capable of specifically binding
the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
proliferation or the activity of a gene in an operon required for
proliferation; (7) identifying a compound that influences the activity of
the gene product or that has an activity against a biological pathway
required for proliferation, or that inhibits cellular proliferation; (8)
identifying a gene required for cellular proliferation or the biological
pathway in which a proliferation-required gene or its gene product lies
or a gene on which the test compound that inhibits proliferation of an
organism acts; (9) manufacturing an antibiotic; (10) profiling a
compound's activity; (11) a culture comprising strains in which the gene
product is overexpressed or underexpressed; (12) determining the extent
to which each of the strains is present in a culture or collection of
strains; or (13) identifying the target of a compound that inhibits the
proliferation of an organism. The anti-sense nucleic acids are useful for
identifying proteins or screening for homologous nucleic acids required
for cellular proliferation to isolate candidate molecules for rational
drug discovery programs, or for screening homologous nucleic acids

CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pcc_sequences

Sequence 195 AA;

Query Match 91.3%; Score 935; DB 6; Length 195;

Best Local Similarity 90.3%; Pred. No. 2e-98;

Matches 176; Conservative 6; Mismatches 13; Indels 0; Gaps 0;

QY 1 MRLTTIGQFPAYQUTALIGDLSKVDKQPGDYFTTTSDEHPGKRVFFPMPDFTFV 60

Db 1 MRLTTIGQFPAYELTALIGDLSKVDKQPGDYFTTTSDEHAGKRVFFPMPDFTFV 60

QY 61 CPTETAPSKNDPEFDRDAQILGVSTIDSEFAHQMAQHNDLTLPPPMKSDIKREL 120

Db 61 CPTETATFGKLNDFEEDDAQVIGVSTIDSEFVHFNMRQHELDLNLPPPMKSDIKREL 120

QY 121 AAGVLMADGVADRVTFIVDPNNIEIQFVSATAGSVGRNVDEVLRVLDALQSDDELCAQNRK 180

Db 121 ATGVLMADGVADRVTFIVDPNNIEIQFVSATAGSVGRNVDEVLRVLDALQSDDELCAQNRK 180

QY 181 GDPTLDAGELKASA 195

Db 181 GDPTLDAGELKASA 195

RESULT 8

AAB67761

ID AAB67761 standard; protein; 195 AA.

AC AAB67761;

DT 11-JUN-2001 (first entry)

DE Amino acid sequence of the IO1 antigen.

KM IO1 antigen; paratuberculosis; IO2 antigen; Crohn's disease; vaccine.

OS Mycobacterium avium.

FH Key Location/Qualifiers

FT Misc-difference 192

PN WO200118212-A2.

PD 15-MAR-2001.

PF 06-SEP-2000; 2000WO-GB003430.

PR 06-SEP-1999; 99NO-00004317.

PA (VESO-) VESO AS.

PI (JONE/) JONES E J.

DR WPI; 2001-244574/25.

DR N-PSDB; AAF80135.

PS Claim 1; Fig 1; 61pp; English.

The present sequence represents the IO1 antigen of *Mycobacterium avium*.
The specification describes a method for diagnosing and monitoring
paratuberculosis infection in a human or non-human animal. The method

CC comprises assessing the presence or concentration of a polypeptide
 CC derived from paratuberculosis, a nucleic acid encoding the polypeptide,
 CC antibodies directed to the polypeptide or one or more infection markers
 CC which result from in vitro presentation of the polypeptide in a sample
 CC from the human or non-human animal. The polypeptide is especially the 101
 CC or 102 antigen. 101 and 102 polypeptide, polynucleotides, and antibodies
 CC are useful for assessing the presence or concentration of
 CC paratuberculosis in a sample. They are useful for stimulating an immune
 CC response against or for preventing infection by *M. avium* ssp
 CC paratuberculosis, for diagnosing and monitoring *M. avium* ssp
 CC paratuberculosis infection and for preventing paratuberculosis or Crohn's
 CC disease in a human or non-human animal, especially a human or a cow
 XX
 SQ Sequence 195 AA;
 Query Match 90.4%; Score 926; DB 4; Length 195;
 Best Local Similarity 89.2%; Pred. No. 2.1e-97;
 Matches 174; Conservative 7; Mismatches 14; Indels 0; Gaps 0;
 QY 1 MFLTTIGDQFPAYQLTALIGDLSKVDKQPGDYFTTITSDHGGKRWVFPMKDTFTFV 60
 DB 1 MFLTTIGDQFPAYQLTALIGDLSKVDKQPGDYFTTITSDHGGKRWVFPMKDTFTFV 60
 QY 61 CPTETIAAFSKLNDSEFEDRDAGIIGVSIIDSEFAHFQMRQANDLKTLPFPMLSDIKRELSQ 120
 DB 61 CPTETIAAFSKLNDSEFEDRDAGIIGVSIIDSEFAHFQMRQANDLKTLPFPMLSDIKRELSQ 120
 QY 121 AAGVLTNDGVADRVTFIVDPNNEIQFVSATAGSVGRNVDEVLRVLDALQSDDELCAQMRK 180
 DB 121 AAGVLTNDGVADRVTFIVDPNNEIQFVSATAGSVGRNVDEVLRVLDALQSDDELCAQMRK 180
 QY 181 GDPTLDAGEELUKASA 195
 DB 181 GDPTLDAGEELUKASA 195
 QY 181 GDPTLDAGEELUKASA 195
 DB 181 GDPTLDAGEELUKASA 195
 RESULT 9
 ABU36067
 ID ABU36067 standard; protein; 195 AA.
 XX
 AC ABU36067;
 XX
 DT 19-JUN-2003 (first entry)
 XX
 DE Protein encoded by Prokaryotic essential gene #21594.
 XX
 KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
 OS
 XX Mycobacterium leprae.
 OS
 PN WO200277183-A2.
 XX
 PD 03-OCT-2002.
 XX
 PF 21-MAR-2002; 2002WO-US009107.
 XX
 PR 21-MAR-2001; 2001US-00815242.
 XX
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX
 PI Mang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW,
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX
 DR WPI; 2003-029926/02.
 DR N-PSDB; ACA39937.
 XX
 PT New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.

XX
 PS Claim 25; SEQ ID NO 63991; 1766pp; English.
 XX
 CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 195 AA;
 Query Match 88.3%; Score 904; DB 6; Length 195;
 Best Local Similarity 86.7%; Pred. No. 7.2e-95;
 Matches 169; Conservative 12; Mismatches 14; Indels 0; Gaps 0;
 QY 1 MFLTTIGDQFPAYQLTALIGDLSKVDKQPGDYFTTITSDHGGKRWVFPMKDTFTFV 60
 DB 1 MFLTTIGDQFPAYQLTALIGDLSKVDKQPGDYFTTITSDHGGKRWVFPMKDTFTFV 60
 QY 61 CPTETIAAFSKLNDSEFEDRDAGIIGVSIIDSEFAHFQMRQANDLKTLPFPMLSDIKRELSQ 120
 DB 61 CPTETIAAFSKLNDSEFEDRDAGIIGVSIIDSEFAHFQMRQANDLKTLPFPMLSDIKRELSQ 120
 QY 121 AAGVLTNDGVADRVTFIVDPNNEIQFVSATAGSVGRNVDEVLRVLDALQSDDELCAQMRK 180
 DB 121 AAGVLTNDGVADRVTFIVDPNNEIQFVSATAGSVGRNVDEVLRVLDALQSDDELCAQMRK 180
 QY 181 GDPTLDAGEELUKASA 195
 DB 181 GDPTLDAGEELUKASA 195
 QY 181 GDPTLDAGEELUKASA 195
 DB 181 GDPTLDAGEELUKASA 195
 RESULT 10
 AAR14303
 ID AAR14303 standard; protein; 193 AA.
 XX
 AC AAR14303;
 XX
 DT 02-JAN-1992 (first entry)
 XX
 DE *M. avium* av1-3 protein.
 XX
 DR mycobacterial infection; diagnostic kit.
 KW
 XX Mycobacterium avium.
 OS
 XX JP03206098-A.

XX 09-SEP-1991.
PD 29-DEC-1989; 89JP-00344183.
XX 29-DEC-1989; 89JP-00344183.
XX 29-DEC-1989; 89JP-00344183.
XX (AJIN) AJINOMOTO KK.
XX WPI; 1991-307344/42.
DR N-PSDB; AAQ14175.
XX Mycobacterium avium derived AVI-3 protein - prepd. from new transformant
PT transformed by plasmid, used as diagnostic agent for mycobacterium avium
XX infection.
XX Claim 1; Page 1 and Fig 3; 12pp; Japanese.
XX The sequence coding for the avi-3 protein is also claimed. Transformants
CC containing the coding sequence are cultured to produce recombinant avi-3
CC for diagnosis of M.avium infection
XX SQ . Sequence 193 AA;

Query Match 88.0%; Score 901.5; DB 2; Length 193;
Best Local Similarity 88.1%; Pred. No. 1.4e-94;
Matches 171; Conservative 6; Mismatches 16; Indels 1; Gaps 1;

QY 2 PLUTTGDPFAYQVLTALIGDLSKVDKOPGDYFTTTSDEHFGKRVVFPMPKDFTV 61
DB 1 PLUTTGDPFAYQVLTALIGDLSKVDKOPGDYFTTTSDEHFGKRVVFPMPKDFTV 60
QY 62 PTEIAAFSKLNDEFEDRDQAQILGVSIDSEFAHQWRAQHNDLKTLPPMLSDIKRELSQA 121
DB 61 PTEIAAFSKLNDEFEDRDQAQILGVSIDSEFAHQWRAQHNDLKTLPPMLSDIKRELSQA 120
QY 122 AGVLTADGVADRVFTFVDPNNIEQFVSATAGSVGRVDETVRLVDALQSDDELCAQWNRK 181
DB 121 TGVLTADGVADRVFTFVDPNNIEQFVSATAGSVGRVDETVRLVDALQSDDELCAQWNRK 179
QY 182 DPTLTADGELIKASA 195
DB 180 DPTLTADGELIKASA 193

RESULT 11
ADB74335
ID ADB74335 standard; protein; 195 AA.

XX ADB74335;
XX 04-DEC-2003 (first entry)
XX Mycobacterium leprae non-naturally occurring peptide #58.
XX Non-naturally occurring peptide; anion pump protein; tuberculosis;
KW hypersensitivity reaction; tuberculostatic.
XX Mycobacterium leprae.
XX US6583266-B1.
XX 24-JUN-2003.
XX 16-SEP-1994; 94US-00311731.
XX 19-AUG-1993; 93US-00109181.
XX 22-OCT-1993; 93US-00142558.
XX (GENO-) GENOME THERAPEUTICS CORP.
XX Smith DR, Mao J;
XX

DR WPI; 2003-655441/62.
DR N-PSDB; ADB74276.
XX New Mycobacterium tuberculosis anion pump peptide useful for as
PT tuberculosis vaccine and diagnosis of tuberculosis infection.
XX Disclosure; SEQ ID NO 84; 26pp; English.

XX The invention relates to a non-naturally occurring peptide of
CC Mycobacterium tuberculosis comprising an amino acid sequence
CC corresponding to an anion pump protein. The invention also relates to a
CC non-naturally occurring nucleic acid corresponding to a DNA sequence of
CC Mycobacterium tuberculosis or Mycobacterium leprae. The new peptide is
CC useful as a vaccine against Mycobacterium tuberculosis or Mycobacterium
CC leprae or for screening for new tuberculosis drugs. Purified proteins
CC derived from the sequences of the invention may elicit a specific immune
CC response. The peptide may also be used to detect hypersensitivity
CC reactions of individuals exposed to Mycobacterium tuberculosis or
CC Mycobacterium leprae. The proteins and peptides may be affixed to solid
CC supports to detect antibodies typical of hypersensitivity reactions, from
CC a patient's sera. This sequence represents Mycobacterium leprae non-
CC naturally occurring peptide of the invention. Note: The sequence data for
CC this patent did not form part of the printed specification but was
CC obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html.

Sequence 195 AA;

Query Match 87.7%; Score 898; DB 7; Length 195;
Best Local Similarity 85.6%; Pred. No. 3.5e-94;
Matches 167; Conservative 14; Mismatches 14; Indels 0; Gaps 0;

QY 1 MPLTTIGDQFPAVQVLTALIGDLSKVDKOPGDYFTTTSDEHFGKRVVFPMPKDFTV 60
DB 1 MSLLTIGDQFPAVQVLTALIGDLSKVDKOPGDYFTTTSDEHFGKRVVFPMPKDFTV 60
QY 61 CPTTIAAFSKLNDEFEDRDQAQILGVSIDSEFAHQWRAQHNDLKTLPPMLSDIKRELSQ 120
DB 61 CPTTIAAFSKLNDEFEDRDQAQILGVSIDSEFAHQWRAQHNDLKTLPPMLSDIKRELSQ 120
QY 121 AAGVLTADGVADRVFTFVDPNNIEQFVSATAGSVGRVDETVRLVDALQSDDELCAQWNRK 180
DB 121 AAGVLTADGVADRVFTFVDPNNIEQFVSATAGSVGRVDETVRLVDALQSDDELCAQWNRK 180
QY 181 GPTLTADGELIKASA 195
DB 181 GPTLTADGELIKASA 195

RESULT 12
ABU25828
ID ABU25828 standard; protein; 198 AA.

XX ABU25828;
XX 19-JUN-2003 (first entry)
XX Protein encoded by Prokaryotic essential gene #11355.
XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX Corynebacterium diphtheriae.
XX WO200277183-A2.
XX 03-OCT-2002.
XX 21-MAR-2002; 2002WO-US009107.
XX 21-MAR-2001; 2001US-00815242.
XX 06-SEP-2001; 2001US-00948993.
XX 25-OCT-2001; 2001US-0342923P.
XX 08-FEB-2002; 2002US-00072851.
XX

PR 06-MAR-2002; 2002US-0362699P.
XX (ELIT-) ELITRA PHARM INC.
XX
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohleen KL, Zyskind JM;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX WPI; 2003-029926/02.
DR N-PSDB; ACA29698.
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
XX Claim 25; SEQ ID NO 53752; 1766bp; English.
XX
XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pcr_sequences
XX
SQ Sequence 198 AA;

Query Match 67.7%; Score 693; DB 6; Length 198;
Best Local Similarity 67.9%; Pred. No. 1.2e-70;
Matches 125; Conservative 29; Mismatches 30; Indels 0; Gaps 0;
QY 1 MPLLITDQPPAYQVLTALIGDLSKVDAKQPGDYFTTITSDEHKGKRVVFPWKOTFTV 60
DB 1 MSILTVKEKPEEFNLTKGGDLHDVNASQEDYFEVSLDKYGGKKRVVFPFKDFTFV 60
QY CPTETIAAFSKLNDSEFDRDAQILGVSIIDSEFAHFQWRQOHNDLKTLPFPMUSDIGRELSQA 120
DB 61 CPTETIAAFGLNDEDFQDRDQILGSGSIDNESHFNMAKTHBELKTVPFPLSDIKHDLIK 120
QY 121 AAGVLNADGVADRVTFTVDPNNEIQFVSATAGSVGRVNDVLTALDLSQDELCAQWVRK 180
DB 121 ALGVENNEGVADRATFTIIDPGIIGFVSVTPDAVGRVNDVLTALDLSQSEVCAQWVRK 180
QY 181 GDPFT 184
DB 181 NDPT 184

RESULT 13
AEB39863

ID AEB39863 standard; protein; 179 AA.
XX
XX AEB39863;
XX
XX 08-SEP-2005 (first entry)
XX
XX L. pneumophila protein SEQ ID NO 4195.
XX
XX detection; infection; Antibacterial; Vaccine.
XX
XX Legionella pneumophila.
XX
XX W02005049642-A2.
XX
XX 02-JUN-2005.
XX
XX 23-SEP-2004; 2004WO-IB003578.
XX
XX 21-NOV-2003; 2003FR-00013687.
XX
XX (INSP) INST PASTEUR.
XX (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
XX (UYLY-) UNIV LYON 1 BERNARD CLAUDE.
XX (CNRS) CNRS CENT NAT RECH SCI.
XX Buchrieser C, Tichit M, Etienne J, Ma L, Cazale C, Glaeser P,
PI Ruenloek C, Bouchier C, Zidane N, Magnier A, Kunat F, Vandenech F,
PI Jarraud S;
XX WPI; 2005-368305/40.
XX
XX New genome of Legionella pneumophila Paris strain and derived
PT polypeptides, useful for detection or identification of the strain and
PT for treatment and prevention of infections.
XX
XX Claim 3; SEQ ID NO 4195; 660bp; English.
XX
XX The invention relates to an isolated or purified nucleotide sequences (I)
CC from Legionella pneumophila Paris strain. (I), and their related
CC sequences or fragments, are useful as primers and probes for detection
CC and amplification, including differentiation between the Paris and
CC Philadelphia strains of Legionella pneumophila and to prepare recombinant
CC (hybrid) polypeptides (II). (II) are also useful for preparation of
CC specific antibodies (Ab), also used for detection/identification of
CC Legionella, and some (II), specifically those involved in synthesis of
CC surface proteins, are targets for identification of inhibitors. (II), or
CC vectors that contain (II), are useful as vaccines and immunogens
CC compositions, for treatment and prevention of infections by L.
CC pneumophila. The present sequence represents the amino acid sequence of a
CC L. pneumophila protein.
XX
SQ Sequence 179 AA;

Query Match 55.3%; Score 566.5; DB 9; Length 179;
Best Local Similarity 58.2%; Pred. No. 3.3e-56;
Matches 107; Conservative 26; Mismatches 42; Indels 9; Gaps 2;
QY 3 LTTIGDQPPAYQVLTALIGDLSKVDAKQPGDYFTTITSDEHKGKRVVFPWKOTFTVCP 62
DB 1 MITVGKPEPEFOLKATVANDINNA-----FQMSNENYRKXMLVFPWPDFTFVCP 52
QY TETIAAFSKLNDSEFDRDAQILGVSIIDSEFAHFQWRQOHNDLKTLPFPMUSDIGRELSQA 122
DB 63 TETIAAFGRNLNDEPADDAQILGSGTSEFVHLMRWQHPLKDLPPFMLADVGRBELTANL 112
QY 123 GVL-NADGVADRVTFTVDPNNEIQFVSATAGSVGRVNDVLTALDLSQDELCAQWVRK 181
DB 113 GLNDEBGAQGRATFTVDPGQITRFVWVNDLVNGRNPAEVLTVLADLSQDELCPQWVRK 172
QY 182 DPTL 185
DB 173 ERTI 176

```

RESULT 14
ID AEB36453 standard; protein, 198 AA.
XX
AC AEB36453;
XX
DT 08-SEP-2005 (first entry)
XX
DE L. pneumophila protein SEQ ID NO 785.
XX
XX detection; infection; Antibacterial; Vaccine.
XX
OS Legionella pneumophila.
XX
PN MO20050496642-A2.
XX
PD 02-JUN-2005.
XX
PE 23-SEP-2004; 2004MO-IB003578.
XX
PR 21-NOV-2003; 2003PR-00013687.
XX
XX (INSP ) INST PASTEUR.
XX PA (IRRM ) INSERM INST NAT SANTE & RECH MEDICALE.
XX PA (UYLY-) UNIV LYON 1 BERNARD CLAUDE.
XX PA (CNRS ) CNRS CENT NAT RECH SCI.
XX
P1 Buchrieser C, Tichit M, Etienne J, Ma L, Cazalec C, Glaeser P;
P1 Rueschok C, Bouchier C, Zidane N, Magnier A, Kunst F, Vandenesch F;
P1 Jarraud S;
XX
DR WPI; 2005-388305/40.
XX
PT New genome of Legionella pneumophila Paris strain and derived
PT polypeptides, useful for detection or identification of the strain and
PT for treatment and prevention of infections.
XX
PS Claim 3; SEQ ID NO 785; 660pp; English.
XX
XX
XX The invention relates to an isolated or purified nucleotide sequences (I)
XX from Legionella pneumophila Paris strain. (I), and their related
XX sequences or fragments, are useful as primers and probes for detection
XX and amplification, including differentiation between the Paris and
XX Philadelphia strains of Legionella pneumophila and to prepare recombinant
XX (hybrid) polypeptides (II). (II) are also useful for preparation of
XX specific antibodies (Ab), also used for detection/identification of
XX Legionella, and some (II), specifically those involved in synthesis of
XX surface proteins, are targets for identification of inhibitors. (II), or
XX vectors that contain (II), are useful as vaccines and immunogenic
XX compositions, for treatment and prevention of infections by L.
XX pneumophila. The present sequence represents the amino acid sequence of a
XX L. pneumophila protein.
XX
SQ Sequence 198 AA;
XX
Query Match: 55.3%; Score 566.5; DB 9; Length 198;
Best Local Similarity 58.2%; Pred. No. 4e-56;
Matches 107; Conservative 26; Mismatches 42; Indels 9; Gaps 2
3 LTTIGDQPPAYOLTLIGDLSKVAKQPGDYFTTITSDHFGKMRVVFPPMDFTFVCP 62
:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
20 MITVNGKFPFEPDLKATVANDINNA-----FQMLSNFTYNGKMLVVFPPMDFTFVCP 71
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
63 TEIAAFSKINDFEEDRDAQILGVSIDSEFAHFGMRAQNDKTLPPPMISDIRBELSQAA 122
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
72 TEIAEFGRLNSEFAIRDAQILGSGTSEVFNLAHMQHDPDLKDLPPPMADYRRELTANTL 131
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
123 GVL-NADGADVDTFIVDPNNELQFASATGASVGRVNDVLTAVLDLQSDLCACMKWGK 181
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
db 132 GLIDDEEGVAQRAFTFIVDPQGITRFVMTVDLNVGRPAEVLRYVDLQTDLCPCKMKG 191
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
182 DPTL 185

```

DB 192 ERTI 195

RESULT 15
ID ADS30355
ID ADS30355 standard; protein; 188 AA.
XX ADS30355;
AC
XX
DT 02-DEC-2004 (first entry)
XX
DE Bacterial polypeptide #19388.
XX
XX Recombinant DNA construct; transformed plant; improved plant property;
XX cold tolerance; heat tolerance; drought tolerance; herbicide; osmotic;
XX pathogen tolerance; pest tolerance; plant disease resistance;
XX cell cycle pathway modification; plant growth regulator;
XX homologous recombination; seed oil yield; protein yield; carbohydrate;
XX nitrogen; phosphorus; photosynthesis; lignin; galactomanan;
XX bacterial polypeptide.
XX
XX Bacteria.
XX
XX
XX US200323675-A1.
XX
XX
XX 18-DEC-2003.
XX
XX 20-FEB-2003; 2003US-00369493.
XX
XX 21-FEB-2002; 2002US-0360039P.
XX
XX
XX (CAOY/) CAO Y.
XX (HINK/) HINKLE G J.
XX (SLAT/) SLATER S C.
XX (CHEN/) CHEN X.
XX (GOLD/) GOLDMAN B S.
XX
XX
XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX
XX WPI: 2004-061375/06.
XX
XX New recombinant DNA construct comprising a promoter positioned to provide
XX for expression of a polynucleotide encoding a polypeptide from a
XX microbial source, useful for producing plants with improved properties.
XX
XX
XX Claim 1; SEQ ID NO 19388; 122pp; English.

SQ Sequence 188 AA;

Query Match 51.8%; Score 530.5; DB 8; Length 188;

Best Local Similarity 54.6%; Pred. No. 5.1e-52;

Matches 101; Conservative 25; Mismatches 50; Indels 9; Gaps 2;

```
QY      3 LITIGDQFPAYQLTLIGDLSKVDAKQPGDYFTTITSDSHGKKRVVFPWPKDFTFYCP 62
      :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db      5 MLTVGDKIPNFKVKATV-----SLEKGEFODITNETFGKWLVLFPWPKDFTFICP 56
      :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
QY      63 TEIAFSKLNDBFEDRDAQILGVSIDSEFAHPQWRACHNDLKTLPFPMLSDIKRELQAA 122
      |||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db      57 TEIAFGKKKQKQFTDRDAQVLGLSTDSEFVHHAMRTHHPDLNLPFPMADIKHELNAL 116
      |||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
QY      123 GVLN-ADGVADRVTFIVDPNNEIQFVSATAGSVGRNDEVLRVLDALQSDIELCAQNRKG 181
      |:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db      117 GILHKEGVALRATFIADPEGIRHVTYNDLSVGRNVSETIRTLDALQTDIELCPGCMWTKG 176
      |:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
QY      182 DPTLD 186
      :|:|:|
Db      177 EETLD 181
```

Search completed: March 23, 2006, 05:10:50
Job time : 53.844 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 23, 2006, 05:11:18 ; Search time 7.75813 Seconds
(without alignments)
2418.401 Million cell updates/sec

Title: US-10-617-038-26
Perfect score: 1024
Sequence: 1 MPLLTTGDPFPAVOLITALIG.....CNMRKGDPTLDAGELIKASA 195

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: p1r1.*
2: p1r2.*
3: p1r3.*
4: p1r4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1024	100.0	195	2	S71013 alkyl hydroperoxid
2	1024	100.0	195	2	B70679 alkyl hydroperoxid
3	935	91.3	195	2	A43858 alkyl hydroperoxid
4	904	88.3	195	2	B87164 alkyl hydroperoxid
5	444	43.4	184	2	AH3581 alkyl hydroperoxid
6	361.5	35.3	180	2	B69867 2-cys peroxidorexi
7	360	35.2	199	2	S73193 hypothetical prote
8	351.5	34.3	203	2	A12385 peroxidorexin [imp
9	350.5	34.2	200	2	S76284 hypothetical prote
10	345.5	33.7	183	2	A63983 2-cys peroxidorexi
11	336	32.8	178	2	S29119 hypothetical prote
12	330.5	32.3	181	2	AD1638 2-cys peroxidorexi
13	328.5	32.1	181	2	AD1275 2-cys peroxidorexi
14	325.5	31.8	199	2	F83540 probable alkyl hyd
15	325.5	31.8	200	2	AB0552 probable peroxidase
16	322	31.4	226	2	S43598 meers homolog R075
17	320.5	31.3	199	2	A48513 macrophage 23k str
18	320.5	31.3	200	2	AB0368 probable alkyl hyd
19	319.5	31.2	199	2	I52425 probable thioresox
20	318	31.1	257	2	UC0064 MERS protein - mou
21	317.5	31.0	265	2	T09211 baal protein - spi
22	316.5	30.9	198	2	C64715 alkyl hydroperoxid
23	316.5	30.9	242	2	S49173 hydrothetical prote
24	315	30.8	198	2	I68897 probable thioresox
25	314.5	30.7	198	2	H71801 probable peroxidase
26	314.5	30.7	199	2	A46711 probable thioresox
27	314.5	30.7	210	2	T41413 thiol-specific ant
28	314	30.7	192	2	T41413 thioresoxin peroxi
29	312	30.5	198	2	A57716 thiol-specific ant

30	311	30.4	576	2	T16005 hypothetical prote
31	304	29.7	195	2	JC2258 substrate protein
32	303	29.6	197	2	C84951 alkyl hydroperoxid
33	302.5	29.5	200	2	I51016 proliferation asso
34	302	29.5	200	2	G83204 probable peroxidase
35	300.5	29.3	198	2	B81453 alkyl hydroperoxid
36	294.5	28.8	196	2	A47362 thiol-specific ant
37	294.5	28.8	202	2	C97756 thioresoxin peroxi
38	291.5	28.5	200	2	A71689 thioresoxin peroxi
39	289.5	28.3	273	2	A43662 29k peripheral mem
40	284.5	27.8	233	2	S67947 alkyl hydroperoxid
41	282.5	27.6	207	2	E82287 antioxidant; Ahpc/
42	281	27.4	195	2	G71492 probable thio-spec
43	279	27.2	188	2	D71314 probable alkyl hyd
44	278.5	27.2	196	2	S69732 hypothetical prote
45	276.5	27.0	187	2	D87610 alkyl hydroperoxid

ALIGNMENTS

RESULT 1
S71013
alkyl hydroperoxidase reductase chain C - Mycobacterium bovis
C/Species: Mycobacterium bovis
C/Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 09-Jul-2004
C/Accession: S71013; S71011
R/Wilson, T.M.; Collins, D.M.
Mol. Microbiol. 19, 1025-1034, 1996
A/Title: ahpc, A gene involved in isoniazid resistance of the Mycobacterium tuberculosis
A/Reference number: S71011; MUID:96249696; PMID:8830260
A/Accession: S71013
A/Status: nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-195 <MTL>
A/Cross-references: UNIPROT:Q57348; UNIPARC:UPI000002FC7F; EMBL:U24083; NID:G1002370; PII
A/Experimental source: ATCC 35723
A/Note: the nucleotide sequence was submitted to the EMBL Data Library, April 1995
A/Accession: S71011
A/Molecule type: DNA
A/Residues: 1-195 <MTM>
A/Cross-references: UNIPARC:UPI000002FC7F; EMBL:U24084; NID:G1002373; PIDN:AA838112.1; P
A/Experimental source: ATCC 35729
C/Genetics:
A/Gene: ahpc
C/Superfamily: alkyl hydroperoxidase reductase C22 protein; alkyl hydroperoxidase C22 prote
F/23-162/Domain: alkyl hydroperoxidase C22 protein homology <C22>
Query Match 100.0%; Score 1024; DB 2; Length 195;
Best Local Similarity 100.0%; Pred. No. 1e-86;
Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MPLTTGDPFPAVOLITALIGDLSKVDKAKOPGDTFTTTSDEHPGKRRVVFPMKXDTFTV 60
DB 1 MPLTTGDPFPAVOLITALIGDLSKVDKAKOPGDTFTTTSDEHPGKRRVVFPMKXDTFTV 60
QY CPPEIAAFSKUNDEFERDNOIIGVSTDSSEFAHROWAHOHNDKTFEFPMLSDIKRELSQ 120
DB 61 CPPEIAAFSKUNDEFERDNOIIGVSTDSSEFAHROWAHOHNDKTFEFPMLSDIKRELSQ 120
QY 121 AAGVLANADGVADRRTFIVDPNNIEIQFVSATAGSVGRNVDEVLRVLDALQSDDELCAQMRK 180
DB 121 AAGVLANADGVADRRTFIVDPNNIEIQFVSATAGSVGRNVDEVLRVLDALQSDDELCAQMRK 180
QY 181 GDPPTLDAGELIKASA 195
DB 181 GDPPTLDAGELIKASA 195
RESULT 2
B70679
alkyl hydroperoxidase reductase - Mycobacterium tuberculosis (strain H37Rv)
N/Alternate names: ahpc protein

C/Species: Mycobacterium tuberculosis
C/Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C/Accession: B70679; S70169
A/Molecule type: DNA
A/Residues: 1-195 <SHE>
A/Cross-references: UNIPARC:UPI000002FC83; EMBL:U18263; NID:G1040852; PIDN:AAA79917.1; P1
R/Cole, S.T.; Broese, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.
R/Connor, R.; Davies, R.; Devlin, K.; Fellwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
R/Andrews, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squires, S.
Nature 393, 537-544 1998
A/Authors: Sgaree, R.; Slaton, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A/Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A/Reference number: A70500; MUID:98295987; PMID:9634230
A/Accession: B70679
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-195 <COL>
A/Cross-references: UNIPROT:Q57348; UNIPARC:UPI000002FC7F; GB:Z81451; GB:AL123456; NID:G
A/Experimental source: strain H37Rv
R/Dereic, V.; Philipp, W.; Dhandayuthapani, S.; Mudd, M.H.; Curcio, R.; Garbe, T.; Heyt
Mol. Microbiol. 17, 889-900, 1995
A/Title: Mycobacterium tuberculosis is a natural mutant with an inactivated oxidative-st
A/Reference number: S70169; MUID:96123431; PMID:8596438
A/Accession: S70169
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-195 <DER>
A/Cross-references: UNIPARC:UPI000002FC7F; EMBL:U16243; NID:G1172077; PIDN:AAC43585.1; F
C/Genetic8
A/Gene: ampC
C/Superfamily: alkyl hydroperoxide reductase C22 protein; alkyl hydroperoxidase c22 prot
F/23-162/Domain: alkyl hydroperoxidase c22 protein homology <C22>

Query Match 100.0%; Score 1024; DB 2; Length 195;
Best Local Similarity 100.0%; Pred. No. 1e-86;
Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MPLLITGDGPAYQTLALIGDLSKVDKOPGDYFTTITSDEHKGKRVVFWPKDFTFV 60
Db 1 MPLLITGDGPAYQTLALIGDLSKVDKOPGDYFTTITSDEHKGKRVVFWPKDFTFV 60
Qy 61 CPTEIAAFSKLNDSEFDRDAQILGVSIDSEFAHFQWRQHNLDKTLPPMLSDIKREL 120
Db 61 CPTEIAAFSKLNDSEFDRDAQILGVSIDSEFAHFQWRQHNLDKTLPPMLSDIKREL 120
Qy 121 AAGVLNADGVADRVTFIVDPNNEIQFVSATAGSVGRVDEVLRVLDALQSDLCACNWRK 180
Db 121 AAGVLNADGVADRVTFIVDPNNEIQFVSATAGSVGRVDEVLRVLDALQSDLCACNWRK 180
Qy 181 GDPTLDAGELLKASA 195
Db 181 GDPTLDAGELLKASA 195

RESULT 3
A43858
alkyl hydroperoxidase C (EC 1.6.4.-) - Mycobacterium avium
N/Alternate names: Avi-3 antigen
C/Species: Mycobacterium avium
C/Date: 10-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C/Accession: A43858; T09655
R/Yamauchi, R.; Matsumoto, K.; Yamazaki, A.; Takahashi, M.; Fukasawa, Y.; Wada, M.; Abe, C
Infect. Immun. 60, 1210-1216, 1992
A/Title: Cloning and expression of the gene for the Avi-3 antigen of Mycobacterium avium
A/Reference number: A43858; MUID:92175967; PMID:1371765
A/Accession: A43858
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-195 <YAM>
A/Cross-references: UNIPROT:Q57413; UNIPARC:UPI000002FC83; GB:U18263; NID:G1040852; PIDN
A/Note: sequence extracted from NCBI backbone (NCBI:95445, NCBI:95446)
R/Sherran, D.R.; Sabo, P.J.; Hickey, M.J.; Arain, T.M.; Mahatras, G.G.; Yuan, Y.; Barry,
Proc. Natl. Acad. Sci. U.S.A. 92, 6625-6629, 1995
A/Title: Disparate responses to oxidative stress in saprophytic and pathogenic mycobacte
A/Reference number: Z16801; MUID:95327698; PMID:7604044
A/Accession: T09655

A/Status: preliminary; translated from GB/EMBL/DBD
A/Molecule type: DNA
A/Residues: 1-195 <SHE>
A/Cross-references: UNIPARC:UPI000002FC83; EMBL:U18263; NID:G1040852; PIDN:AAA79917.1; P1
C/Genetics:
A/Gene: ampC
C/Superfamily: alkyl hydroperoxide reductase C22 protein; alkyl hydroperoxidase c22 prote
C/Keywords: oxidoreductase
F/23-162/Domain: alkyl hydroperoxidase c22 protein homology <C22>

Query Match 91.3%; Score 935; DB 2; Length 195;
Best Local Similarity 90.3%; Pred. No. 1.6e-78;
Matches 176; Conservative 6; Mismatches 13; Indels 0; Gaps 0;

Qy 1 MPLLITGDGPAYQTLALIGDLSKVDKOPGDYFTTITSDEHKGKRVVFWPKDFTFV 60
Db 1 MPLLITGDGPAYQTLALIGDLSKVDKOPGDYFTTITSDEHKGKRVVFWPKDFTFV 60
Qy 61 CPTEIAAFSKLNDSEFDRDAQILGVSIDSEFAHFQWRQHNLDKTLPPMLSDIKREL 120
Db 61 CPTEIAAFSKLNDSEFDRDAQILGVSIDSEFAHFQWRQHNLDKTLPPMLSDIKREL 120
Qy 121 AAGVLNADGVADRVTFIVDPNNEIQFVSATAGSVGRVDEVLRVLDALQSDLCACNWRK 180
Db 121 AAGVLNADGVADRVTFIVDPNNEIQFVSATAGSVGRVDEVLRVLDALQSDLCACNWRK 180
Qy 181 GDPTLDAGELLKASA 195
Db 181 GDPTLDAGELLKASA 195

RESULT 4
E87164
alkyl hydroperoxide reductase [imported] - Mycobacterium leprae
C/Species: Mycobacterium leprae
C/Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C/Accession: E87164
R/Cole, S.T.; Eigmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Hor
R.; Davies, R.M.; Devlin, K.; Dutuoy, S.; Fellwell, T.; Frazer, A.; Hamlin, N.; Holroyd,
Nature 409, 1007-1011, 2001
A/Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squires, R.; Sq
A/Title: Massive gene decay in the leprosy bacillus.
A/Reference number: A86909; MUID:21128732; PMID:11234002
A/Accession: E87164
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-195 <STO>
A/Cross-references: UNIPROT:Q9CBF5; UNIPARC:UPI00000662C; GB:AL450380; NID:G13093656; P
C/Genetics:
A/Gene: ampC
C/Superfamily: alkyl hydroperoxide reductase C22 protein; alkyl hydroperoxidase c22 prote

Query Match 88.3%; Score 904; DB 2; Length 195;
Best Local Similarity 86.7%; Pred. No. 1.1e-75;
Matches 169; Conservative 12; Mismatches 14; Indels 0; Gaps 0;

Qy 1 MPLLITGDGPAYQTLALIGDLSKVDKOPGDYFTTITSDEHKGKRVVFWPKDFTFV 60
Db 1 MPLLITGDGPAYQTLALIGDLSKVDKOPGDYFTTITSDEHKGKRVVFWPKDFTFV 60
Qy 61 CPTEIAAFSKLNDSEFDRDAQILGVSIDSEFAHFQWRQHNLDKTLPPMLSDIKREL 120
Db 61 CPTEIAAFSKLNDSEFDRDAQILGVSIDSEFAHFQWRQHNLDKTLPPMLSDIKREL 120
Qy 121 AAGVLNADGVADRVTFIVDPNNEIQFVSATAGSVGRVDEVLRVLDALQSDLCACNWRK 180
Db 121 AAGVLNADGVADRVTFIVDPNNEIQFVSATAGSVGRVDEVLRVLDALQSDLCACNWRK 180
Qy 181 GDPTLDAGELLKASA 195
Db 181 GDPTLDAGELLKASA 195

A>Note: Nostrac sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C>Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C/Accession: A12385
R/Kaneke, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasaoto, S.; Matanabe, A.; Iriyuchi, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A>Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena
A/Reference number: AB1807; MUID:21595285; PMID:11759840
A/Accession: A12385
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-203 <KUR>
A/Cross-references: UNIPROT:Q8VNC5; UNIPARC:UPI00000CEBD4; GB:BA000019; PIDD:BA076340.1;
A/Experimental source: strain PCC 7120
C/Genetics:
C/Superfamily: alkyl hydroperoxide reductase C22 protein; alkyl hydroperoxidase c22 protein
Query Match 34.3%; Score 351.5; DB 2; Length 203;
Best Local Similarity 40.9%; Pred. No. 6.7e-25;
Matches 79; Conservative 34; Mismatches 55; Indels 25; Gaps 5;
Qy 4 LTIGDPPAYQLTALIGDLSKVDAKQPGDYFTTITSDHFGKRVVFPKDFTCPT 63
Db 11 LRVGQAPDPTATAVVQD-----FKTKLSDYRGKTVLFFYPLDTFTVCPT 58
Qy 64 EIAAFSKLNDEFDRDAQILGVSIDSEFAHFQWRAQNDLKT-----LPPMLSDIKREL 118
Db 59 EITAFSRYEERFKKLTIELIGSVDSSEFSLAW--IQDRKSGGVGLNLYLVSIDIKREV 116
Qy 119 SQAAGVLA-ADGVADRFTFYDPPNNEIOFVSATAGSVGRNDEVLRVDAQ-----SDE 172
Db 117 SDATNVLDPAAGIALRGFIIDKGIQIQTATINNLAFGRSVDETLFTLQALQVQSHPDE 176
Qy 173 LCACNMRKGDPTL 185
Db 177 VCPAGMQPGKMTM 189
RESULT 9
576284
hypothetical protein 3 - Synechocystis sp. (strain PCC 6803)
C/Species: Synechocystis sp.
A/Variety: PCC 6803
C/Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C/Accession: S76284
R/Kaneke, T.; Sato, S.; Korani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.; O, K.; Okumura, S.; Shimo, S.; Takeuchi, C.; Wada, T.; Matanabe, A.; Yamada, M.; Yasuda, M.
DNA Res. 3, 109-136, 1996
A>Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
A/Reference number: S74322; MUID:97061201; PMID:8905231
A/Accession: S76284
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-200 <KAN>
A/Cross-references: UNIPROT:Q55624; UNIPARC:UPI0000139C78; EMBL:D64000; GB:AB001339; NID:
A/Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C/Superfamily: alkyl hydroperoxide reductase C22 protein; alkyl hydroperoxidase c22 protein
F;14-155/Domain: alkyl hydroperoxidase c22 protein homology <C22>
Query Match 34.2%; Score 350.5; DB 2; Length 200;
Best Local Similarity 38.5%; Pred. No. 8.2e-25;
Matches 74; Conservative 40; Mismatches 57; Indels 21; Gaps 5;
Qy 3 LVTIGDPPAYQLTALIGDLSKVDAKQPGDYFTTITSDHFGKRVVFPKDFTCPT 62
Db 4 LRVGQAPDPTATAVVQD-----FKTKLSDYRGKTVLFFYPLDTFTVCPT 51
Qy 63 TEIAAFSKLNDEFDRDAQILGVSIDSEFAHFQW---RAQNDLKTLPFPMLSIDIKREL 119
Db 52 TEIAAFSDRHSFALDTEVVGISVDSSEFSLAWIQERKMGKGININYPVSDIKREIS 111

Qy 120 QAAGVLAAD-GVADRFTFYDPPNNEIOFVSATAGSVGRNDEVLRVDAQ-----SDEL 173
Db 112 QAQNVLEPPDAGIALRGFIIDKGIQIQTATINNLAFGRSVDETLFTLQALRHVQSHNEV 171
Qy 174 CACNMRKGDPTL 185
Db 172 CPVDMDQBGDKTM 183
RESULT 10
A83983
2-cys peroxiredoxin BH265 [imported] - Bacillus halodurans (strain C-125)
C/Species: Bacillus halodurans
C/Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C/Accession: A83983
R/Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hiran
Nucleic Acids Res. 28, 4317-4331, 2000
A>Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A/Reference number: A83650; MUID:20512582; PMID:11058132
A/Accession: A83983
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-183 <STO>
A/Cross-references: UNIPROT:Q9K912; UNIPARC:UPI00000C3F86; GB:AP001516; GB:BA000004; NID:
A/Experimental source: strain C-125
C/Genetics:
A/Gene: BH265
C/Superfamily: alkyl hydroperoxide reductase C22 protein; alkyl hydroperoxidase c22 protein
Query Match 33.7%; Score 345.5; DB 2; Length 183;
Best Local Similarity 39.2%; Pred. No. 2.1e-24;
Matches 73; Conservative 30; Mismatches 66; Indels 17; Gaps 4;
Qy 6 IGDQPPAYQLTALIGDLSKVDAKQPGDYFTTITSDH--PGKRVVFPKDFTCPT 63
Db 7 VAKQAPRFEM-----DAVMNPKERGVSLSENMKNQDKWTVLFFYPMDFTCPT 55
Qy 64 EIAAFSKLNDEFDRDAQILGVSIDSEFAHFQW---RAQNDLKTLPFPMLSIDIKREL 120
Db 56 EITSLSDRYEFEDLDIAEVIGSTDTIRHKAWINTSRDNGIDGLKPLADATNHEVSR 115
Qy 121 AAGVLA-ADGVADRFTFYDPPNNEIOFVSATAGSVGRNDEVLRVDAQSDSLCAQNR 179
Db 116 EYGVLAIEBGLIALRGFIIDKGIQIQTATINNLAFGRSVDETLFTLQALGCLPAMNK 175
Qy 180 KGDPTL 185
Db 176 PQGETL 181
RESULT 11
S29119
hypothetical protein 3 - Clostridium pasteurianum
C/Species: Clostridium pasteurianum
C/Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
C/Accession: S29119
R/Mathieu, I.; Meyer, J.; Moulis, J.M.
Biochem. J. 285, 255-262, 1992
A>Title: Cloning, sequencing and expression in Escherichia coli of the rubredoxin gene fr
A/Reference number: S29117; MUID:92344580; PMID:1637309
A/Accession: S29119
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-178 <MAT>
A/Cross-references: UNIPROT:P23161; UNIPARC:UPI00002FEC66; EMBL:M60116; NID:9144905; PIDD:
A/Note: alkyl hydroperoxide reductase C22 protein; alkyl hydroperoxidase c22 protein
F;12-155/Domain: alkyl hydroperoxidase c22 protein homology <C22>
Query Match 32.8%; Score 336; DB 2; Length 178;
Best Local Similarity 38.6%; Pred. No. 1.5e-23;
Matches 71; Conservative 34; Mismatches 65; Indels 14; Gaps 3;
Qy 6 IGDQPPAYQLTALIGDLSKVDAKQPGDYFTTITSDHFGKRVVFPKDFTCPT 65

```
Db      5 VGKPAPEFEMKAV-----KGDGKGFTEVKGDKGKMLWFFPDLPTFVCPTEI 54
Qy      66 AAFSKLNDPEFDRDAQILGVSIDSEFAHFQWR---HNDKTLPPFMLSIDIKRELSDQA 122
      55 TGFESKRAEEFDDLKAEILLAVSCDSQYSHETWINDIKOGGIGKINFPILASKTTEVSTKY 114
Qy      123 GV-LNADGVADRVTFTVDPNNEIQFVSATAGSVGRNVDVLRVLDALQSDDELCAQNMKG 181
      115 GIGIEEGSISLRGFIIDPEGIVRYSVVDLNVGRSVDETRVLKAPQTGMCAQLDMHEG 174
Qy      182 DPTL 185
      175 DDNL 178

RESULT 12
AD1638
2-cys peroxiredoxin homolog l1n1645 [imported] - Listeria innocua (strain Clp11262)
C/Species: Listeria innocua
C/Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C/Accession: AD1638
R/Glaeser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker,
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurge, O.; Entian, K.D.; Feihl, H.
D.; Jones, L.M.; Karet, U.
Science 294, 849-852, 2001
A/Authors: Kretz, J.; Simoes, N.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend,
A./Title: Comparative genomics of Listeria species.
A/Reference number: AB1077; MUID:21537279; PMID:11679669
A/Accession: AD1638
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-181 <GLA>
A/Cross-references: UNIPROT:Q92BA5; UNIPARC:UPI00000CC601; GB:AL592022; PIDN:CAC96876.1;
A/Experimental source: strain Clp11262
C/Genetics:
A/Gene: l1n1645
C/Superfamily: alkyl hydroperoxide reductase C22 protein; alkyl hydroperoxidase c22 pro
Query Match      32.3%; Score 330.5; DB 2; Length 181;
Best Local Similarity 37.1%; Pred. No. 4.9e-23;
Matches 69; Conservative 38; Mismatches 62; Indels 17; Gaps 4;

Qy      6 IGDPFAYVLTALIGDLSKYDAKQPGDYFTTTSDEH--PGKRVVFEMFKDFTFVCPT 63
      6 VGTQAPFEFEMKAVM-----PNOTGKVSLEKNIEDDKMTVLFYPMDFVCPTEI 54
Qy      64 EIAAFSKLNDPEFDRDAQILGVSIDSEFAHFQWR---AQHNDKTLPPFMLSIDIKRELSDQ 120
      55 EIVAIASRSDEFDLAKRIITGASTDTTSHLAWNTPTIKGGIGKINFPILADTNHGVAS 114
Db      121 AAGVL-NADGVADRVTFTVDPNNEIQFVSATAGSVGRNVDVLRVLDALQSDDELCAQNMKG 179
      115 DYGVLIEEGVALRGLFTINPKGEIQEVVHNHNGREVDVLRVLDALQSGCLPIMWQ 174
Qy      180 KGDPPTL 185
      175 PGERTI 180

RESULT 13
AD1275
2-cys peroxiredoxin homolog lmo1604 [imported] - Listeria monocytogenes (strain EGD-e)
C/Species: Listeria monocytogenes
C/Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C/Accession: AD1275
R/Glaeser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker,
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurge, O.; Entian, K.D.; Feihl, H.
D.; Jones, L.M.; Karet, U.
Science 294, 849-852, 2001
A/Authors: Kretz, J.; Kunst, M.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend,
```

```
A/Title: Comparative genomics of Listeria species.
A/Reference number: AB1077; MUID:21537279; PMID:11679669
A/Accession: AD1275
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-181 <GLA>
A/Cross-references: UNIPROT:Q816S9; UNIPARC:UPI00000CF186; GB:NC_003210; PIDN:CAC99682.1;
A/Experimental source: strain EGD-e
C/Genetics:
A/Gene: lmo1604
C/Superfamily: alkyl hydroperoxide reductase C22 protein; alkyl hydroperoxidase c22 pro
Query Match      32.1%; Score 328.5; DB 2; Length 181;
Best Local Similarity 36.6%; Pred. No. 7.5e-23;
Matches 68; Conservative 39; Mismatches 62; Indels 17; Gaps 4;

Qy      6 IGDPFAYVLTALIGDLSKYDAKQPGDYFTTTSDEH--PGKRVVFEMFKDFTFVCPT 63
      6 VGTQAPFEFEMKAVM-----PNOTGKVSLEKNIEDDKMTVLFYPMDFVCPTEI 54
Qy      64 EIAAFSKLNDPEFDRDAQILGVSIDSEFAHFQWR---AQHNDKTLPPFMLSIDIKRELSDQ 120
      55 EIVAIASRSDEFDLAKRIITGASTDTTSHLAWNTPTIKGGIGKINFPILADTNHGVAS 114
Db      121 AAGVL-NADGVADRVTFTVDPNNEIQFVSATAGSVGRNVDVLRVLDALQSDDELCAQNMKG 179
      115 DYGVLIEEGVALRGLFTINPKGEIQEVVHNHNGREVDVLRVLDALQSGCLPIMWQ 174
Qy      180 KGDPPTL 185
      175 PGERTI 180

RESULT 14
F83540
probable alkyl hydroperoxide reductase PA0848 [imported] - Pseudomonas aeruginosa (strain
C/Species: Pseudomonas aeruginosa
C/Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C/Accession: F83540
R/Stover, C.K.; Pham, X.O.; Erwin, A.L.; Mitsuuchi, S.D.; Warren, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardis, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A/Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A/Reference number: AB2950; MUID:20437337; PMID:10984043
A/Accession: F83540
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-199 <STO>
A/Cross-references: UNIPROT:Q91593; UNIPARC:UPI00000C5187; GB:AE004519; GB:AE004091; NID:
A/Experimental source: strain PA01
C/Genetics:
A/Gene: PA0848
C/Superfamily: alkyl hydroperoxide reductase C22 protein; alkyl hydroperoxidase c22 pro
Query Match      31.8%; Score 325.5; DB 2; Length 199;
Best Local Similarity 37.1%; Pred. No. 1.6e-22;
Matches 73; Conservative 39; Mismatches 62; Indels 23; Gaps 5;

Qy      6 IGDPFAYVLTALIGDLSKYDAKQPGDYFTTTSDEH--PGKRVVFEMFKDFTFVCPT 65
      5 VNRQAPDFTAAAVL--GDGSIVDAFQLSSL-----RGKYVVLFPWPIDFVCPSEI 54
Db      66 AAFSKLNDPEFDRDAQILGVSIDSEFAHFQWR---QHNDKTLPPFMLSIDIKRELSDQA 122
      55 IAHNNRDKREIEGVEVGVGVSIDSQFTTHAMRSPTVKKGIGIAVEFVWADVAGHEITRAY 114
Qy      123 GV-LNADGVADRVTFTVDPNNEIQFVSATAGSVGRNVDVLRVLDALQSDDELCAQNMKG 178
      115 GIEHEDVALRASGLIRAGVGHQVNNPLRGREVDVLRVLDALQFTBEHGVCPAGW 174
Db      179 RKGDPTLADGELLKASA 195
      |||
```

Db 175 RKQ-----KGMKSA 185

RESULT 15

AB0552

probable peroxidase [imported] - *Salmonella enterica* subsp. *enterica* serovar Typh (stre

C:Species: *Salmonella enterica* subsp. *enterica* serovar Typh

A:Note: this species has also been called *Salmonella typh*

C>Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002

C/Accession: AB0552

R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Main, J.; Churcher, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moule, S.; O'Gaora, P.

Nature 413, 848-852, 2001

A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;

A>Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* serov

A:Reference number: AB0502; MUID:21534947; PMID:11677608

A:Accession: AB0552

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-200 <PAR>

A:Cross-references: UNIPARC:UP1000005A247; GB:AL513382; PIDN:CAD08858.1; PID:G16501671;

C:Genetics:

A:Gene: STY0440

C:Superfamily: alkyl hydroperoxide reductase C22 protein; alkyl hydroperoxidase c22 prot

Query Match 31.8%; Score 325.5; DB 2; Length 200;

Best Local Similarity 35.6%; Pred. No. 1.6e-22;

Matches 69; Conservative 39; Mismatches 67; Indels 19; Gaps 5;

QY 3 LRTIGDQFPAYQLTALIGDLSKVDAKQPGDYFTTITSDEHP-GKMRVFPFMPKDTFTVC 61

Db 1 MVLVTRQAPDPTAAAVLGS-----GEIVDKENFKQHTNGKTTVLFPFMPDFTFVC 50

QY 62 PTEIAFSKLNDEFEDRDAQILGVISIDSEFAHFQMR--AQHNDLKTLPFPMLSDIKREL 118

Db 51 PSELIAFDKRYEERQKRGVEVGVSPDSEFVHNAMKNTLPVDKSGIGPVKYAMADVGREI 110

QY 119 SQAAGVLNAD-GVADRVTFIVDPNNEIQVSATAGSVGRNVDEVLRVLDALQSDE---L 173

Db 111 OKAYGIEHPDEGVLRGSLIDANGIVRHQVNVNDLPLGRNIDEMLRNVLDALQFHEHGDV 170

QY 174 CACWRRKGDPTLDA 187

Db 171 CPAQWEKKGEGMNA 184

Search completed: March 23, 2006, 05:26:32
Job time : 7.75813 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: March 23, 2006, 04:56:43 ; Search time 49.8737 Seconds
(without alignments)

2758.529 Million cell updates/sec

Title: US-10-617-038-26

Perfect score: 1024
Sequence: 1 MPLLITIGDQPPAYQVLTALIG.....CNMRKDDPTLDAGELKASA 195

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1024	100.0	195	2	Q79CV0_MYCBO
2	1024	100.0	195	2	Q7BHK8_MYCTU
3	943	92.1	195	2	087323_MYCMR
4	935	91.3	195	2	057413_MYCAV
5	934	91.2	195	2	Q73213_MYCPA
6	920	89.8	195	2	Q9AGF6_MYCPA
7	913	89.2	195	2	05YT54_MYCPA
8	904	88.3	195	2	Q9CBF5_MYCLE
9	878	85.7	195	2	057529_MYCSM
10	693	67.7	198	2	Q6NGT3_CORDI
11	685	66.9	198	2	Q46025_CORDI
12	619	60.4	194	2	Q9X5V0_STRVD
13	593	57.9	184	2	Q9FBB5_STRCO
14	592	57.8	184	2	Q9RNT2_STRCO
15	589	57.6	184	2	0821C6_STRAV
16	589	57.5	178	2	Q4NTM9_9DELT
17	569.5	55.6	179	2	Q83BM6_COXBU
18	566.5	55.3	179	2	Q5X2T8_LEGPA
19	565.5	55.2	179	2	Q5WU96_LEGPA
20	564.5	55.1	212	2	Q5ZT15_LEGPH
21	455.5	44.5	180	2	Q5NLQ4_ZYMOA
22	453	44.2	184	2	Q578J3_BRUBA
23	453	44.2	184	2	Q8FVW3_BRUBU
24	444	43.4	184	2	Q8YCF3_BRUME
25	444	43.4	184	2	Q9ANK9_BRALJA
26	429.5	41.9	230	2	Q8XWY0_RALSO
27	423.5	41.4	182	2	Q62J18_BURMA
28	423.5	41.4	182	2	Q63T73_BURPS
29	410.5	40.1	182	2	Q7VTT5_BORPE
30	410.5	40.1	182	2	Q7W5D3_BORPA
31	410.5	40.1	182	2	Q7WCW8_BORBR

32	404	39.5	181	2	Q5QZB4_IDLIO	Q5QZB4_idiomarina
33	389	38.0	179	2	Q67QY2_SYMRH	Q67QY2_symbiobacte
34	374	36.5	180	2	Q74ER7_GEOSL	Q74ER7_geobacter s
35	371	36.2	194	2	Q7V1K9_PROMP	Q7V1K9_prochloroco
36	361.5	35.3	180	2	Q34564_BACSU	Q34564_bacillus su
37	360.5	35.2	180	2	Q65K82_BACID	Q65K82_bacillus li
38	360	35.2	199	1	YCF42_FORPU	YCF42_porphyra pu
39	358.5	35.0	200	2	Q7V7J1_PROMA	Q7V7J1_prochloroco
40	357	34.9	197	2	Q7VBM4_PROMA	Q7VBM4_prochloroco
41	353.5	34.5	200	2	Q7U6X3_SYMPX	Q7U6X3_synechococc
42	352	34.4	197	2	Q8D1X7_SYNEL	Q8D1X7_synechococc
43	351.5	34.3	203	2	Q8YNC5_TANAS	Q8YNC5_anabaena ap
44	351	34.3	199	2	Q7KGY6_LEICH	Q7KGY6_leishmania
45	351	34.3	199	2	Q95NFS_LEITIN	Q95NFS_leishmania

ALIGNMENTS

Q79CV0_MYCBO	PRELIMINARY:	PRT: 195 AA.
ID	Q79CV0_MYCBO	Q79CV0_050557; Q50558; Q50572; Q50573; Q50574; Q57348;
AC	Q79CV0_050557; Q50558; Q50572; Q50573; Q50574; Q57348;	
DT	05-JUL-2004 (TrEMBLrel. 27, Created)	
DT	05-JUL-2004 (TrEMBLrel. 27, Last sequence update)	
DT	13-SEP-2005 (TrEMBLrel. 31, Last annotation update)	
DE	Alkyl hydroperoxide reductase C peptide (ALKYL HYDROPEROXIDE REDUCTASE	
DE	C PROTEIN AHPC (ALKYL HYDROPEROXIDASE C) (EC 1.-.-.-.)	
GN	Name=ahpc; OrderedLocustNames=Mb2454;	
OS	Mycobacterium bovis.	
OC	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;	
OC	Corynebacterineae; Mycobacteriaceae; Mycobacterium;	
OC	Mycobacterium tuberculosis complex.	
OX	NCBI_TaxID=1765;	
RN	[1]	
RP	NUCLEOTIDE SEQUENCE.	
RC	STRAIN=ATCC 35723, and ATCC 35729;	
RA	MEDLINE=96249696; PubMed=8030260;	
RA	Wilson T.M., Collins D.M.;	
RT	"ahpc, a gene involved in isoniazid resistance of the Mycobacterium	
RT	tuberculosis complex."	
RL	Mol. Microbiol. 19:1025-1034(1996).	
RN	[2]	
RP	NUCLEOTIDE SEQUENCE.	
RC	STRAIN=AF2122/97;	
RA	MEDLINE=22709107; PubMed=12788972; DOI=10.1073/pnas.1130426100;	
RA	Garnier T., Eigmeier K., Camus J.-C., Medina N., Mansoor H.,	
RA	Pryor M., Dutfoy S., Grondin S., Lacroix C., Monsemp C., Simon S.,	
RA	Harris B., Altin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,	
RA	Parikhil J., Barrett B.G., Cole S.T., Gordon S.V., Hewinson R.G.;	
RT	"The complete genome sequence of Mycobacterium bovis."	
RL	Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).	
DR	EMBL; U24083; AAB60203.1; -; Genomic DNA.	
DR	EMBL; U24084; AAB38112.1; -; Genomic DNA.	
DR	EMBL; BX248342; CAD97315.1; -; Genomic DNA.	
DR	PIR; B70679; B70679.	
DR	PIR; S71013; S71013.	
DR	SMR; Q79CV0; 2-179.	
DR	GO; GO:0016491; F:oxidoreductase activity; IEA.	
DR	GO; GO:0004601; F:peroxidase activity; IEA.	
DR	InterPro; IPR000866; Ahpc-TSA.	
DR	InterPro; IPR012336; Thiooxidin-like.	
DR	Pfam; PF00578; Ahpc-TSA; 1.	
KM	Complete proteome; Oxidoreductase; Peroxidase.	
SQ	SEQUENCE 195 AA; 21566 MW; 011C101F07C7095 CRC64;	
Q7	Query Match	100.0%; Score 1024; DB 2; Length 195;
Db	Best Local Similarity	100.0%; Pred. No. 8.7e-82;
	Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
	1 MPLLITIGDQPPAYQVLTALIGDLSKVDKAPGQVFTTITSDHGGKRVVFPWPKDFTFV 60	

```
QY 61 CPTBIAAFSKLNDEFEDRDAQILGVSIDSEFAHFQWRAQHNLDKTLPPMLSDIKRELISQ 120
D 61 CPTBIAAFSKLNDEFEDRDAQILGVSIDSEFAHFQWRAQHNLDKTLPPMLSDIKRELISQ 120
QY 121 AAGVLNADGVADRVTVTPVNNETQFVSATAGSVGRNVDEVLRVLDALQSDDELCAQWNRK 180
D 121 AAGVLNADGVADRVTVTPVNNETQFVSATAGSVGRNVDEVLRVLDALQSDDELCAQWNRK 180
QY 181 GDPTLDAGELLKASA 195
D 181 GDPTLDAGELLKASA 195
Db 181 GDPTLDAGELLKASA 195

RESULT 2
QY 07BHK8 MYCTU PRELIMINARY; PRT; 195 AA.
ID 07BHK8 079FE2: 07D758:
AC 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE ALKYL1 HYDROPEROXIDASE C (ALKYL HYDROPEROXIDE REDUCTASE C PROTEIN AHPC
GN Name=ahpc; OrderedLocNames=MT2503, Rv2428;
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinomycetidae; Actinomycetales;
OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium tuberculosis complex.
OX NCBI_Taxid=1773;
RN 1]
RP NUCLEOTIDE SEQUENCE.
RA Orru G., Iona E., Memmi G., Oggioni M.R., Factorini L., Orefici G.,
RA Pozzi G.;
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
RN 2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=H37Rv;
RX MEDLINE=98295987; Pubmed=9634230; DOI=10.1038/31159;
RA Cole S.T., Broesch R., Parhill J., Garner T., Churcher C.M.,
RA Harris D.E., Gordon S.V., Eigmeier K., Gae S., Barry C.E. III,
RA Tekala F., Badcock K., Baaham D., Brown D., Chillingworth T.,
RA Connor R., Davies R.M., Devlin K., Felwell T., Gentles S., Hamlin N.,
RA Holroyd S., Hornsby T., Jęgels K., Krogh A., McLean J., Moule S.,
RA Murphy L.D., Oliver S., Osborne K., Quail M.A., Rajandream M.A.,
RA Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sultson J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544(1998).
RN 3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CDC 1551 / Oshkosh;
RX MEDLINE=22206494; Pubmed=12218036;
RX DOI=10.1128/JB.184.19.5479-5490.2002;
RA Fietechmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J.D., Deboy R.T., Dodson R.J., Gwyn L.D., Haft D.H.,
RA Hickley E.K., Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D.,
RA Salzberg S.L., Delcher A., Uitterback T.R., Feldman J.F., Khouri H.M.,
RA Gill J., Mikula A., Bishai W., Jacobs W.R. Jr., Venter J.C.,
RA Fraser C.M.;
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL J. Bacteriol. 184:5479-5490(2002).
RN 4]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=H37Rv;
RX MEDLINE=96123431; Pubmed=8596438;
RA Dereic V., Philipp W., DhandaYuthapani S., Mudd M.H., Curcio R.,
RA Garbe T., Heym B., Via L.E., Cole S.T.;
RT "Mycobacterium tuberculosis is a natural mutant with an inactivated
RT oxidative-stress regulatory gene: implications for sensitivity to
RT isoniazid.";
RL Mol. Microbiol. 17:889-900(1995).
```

```
RN 5]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=95327698; Pubmed=7604044;
RA Sherman D.R., Sabo P.J., Hickey M.J., Arain T.M., Mahatras G.G.,
RA Yuan Y., Barry C.E. III, Stover C.K.;
RT "Disparate responses to oxidative stress in saprophytic and pathogenic
RT Mycobacteria.";
RL Proc. Natl. Acad. Sci. U.S.A. 92:6625-6629(1995).
DR EMBL: AF313459; AAC34172.1; -; Genomic_DNA.
DR EMBL: BX842579; CAB03768.1; -; Genomic_DNA.
DR EMBL: AF313460; AAC34173.1; -; Genomic_DNA.
DR EMBL: AF000516; AA46800.1; -; Genomic_DNA.
DR EMBL: AF313461; AAC34174.1; -; Genomic_DNA.
DR EMBL: AF313463; AAC34176.1; -; Genomic_DNA.
DR EMBL: U16264; AAA79919.1; -; Genomic_DNA.
DR EMBL: U16243; AAC43585.1; -; Genomic_DNA.
DR EMBL: AF313462; AAC34175.1; -; Genomic_DNA.
DR PDB: 2BMX; X-ray; A/B/C=1-195.
DR SMR; 07BHK8; 2-179.
DR TIGR; MT2503; -.
DR Tuberculist; Rv2428; -.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0004601; F:peroxidase activity; IEA.
DR InterPro; IPR000866; Ahpc-TSA.
DR InterPro; IPR012336; Thioresdoxin-like.
DR InterPro; IPR012335; Thioresdoxin-fold.
DR Pfam; PF00578; Ahpc-TSA; 1.
KW Complete proteome; Oxidoreductase; Peroxidase.
SQ SEQUENCE 195 AA; 21566 MW; 011C1014F07C7095 CRC64;

Query Match 100.0%; Score 1024; DB 2; Length 195;
Best Local Similarity 100.0%; Pred. No. 8; 7e-82;
Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPLLITIGDOPPAQQLATLIGSDLSKVDKOPGDFYTTISDSBPGKRVYFFPKDFTFY 60
D 1 MPLLTIGDOPPAQQLATLIGSDLSKVDKOPGDFYTTISDSBPGKRVYFFPKDFTFY 60
QY 61 CPTBIAAFSKLNDEFEDRDAQILGVSIDSEFAHFQWRAQHNLDKTLPPMLSDIKRELISQ 120
D 61 CPTBIAAFSKLNDEFEDRDAQILGVSIDSEFAHFQWRAQHNLDKTLPPMLSDIKRELISQ 120
QY 121 AAGVLNADGVADRVTVTPVNNETQFVSATAGSVGRNVDEVLRVLDALQSDDELCAQWNRK 180
D 121 AAGVLNADGVADRVTVTPVNNETQFVSATAGSVGRNVDEVLRVLDALQSDDELCAQWNRK 180
QY 181 GDPTLDAGELLKASA 195
D 181 GDPTLDAGELLKASA 195
Db 181 GDPTLDAGELLKASA 195

RESULT 3
QY 087323 MYCWR PRELIMINARY; PRT; 195 AA.
ID 087323 MYCWR PRELIMINARY;
AC 087323;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE Alkylhydroperoxide reductase.
GN Name=ahpc;
OS Mycobacterium marinum.
OC Bacteria; Actinobacteria; Actinomycetidae; Actinomycetales;
OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium;
OX NCBI_Taxid=1781;
RN 1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 15069;
RX MEDLINE=98406038; Pubmed=9733688;
RA Pagan-Ramos E., Song J., McFalone M., Mudd M.H., Dereic V.;
RT "Oxidative stress response and characterization of the oxyR-ahpc and
RT fura-katg loci in Mycobacterium marinum.";
RL J. Bacteriol. 180:4856-4864(1998).
DR EMBL: AF034861; AAC61301.1; -; Genomic_DNA.
```


DR HSP, Q9TX2, 1E2Y.
 DR SMK; O87323; 3-179.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0004601; F:peroxidase activity; IEA.
 DR InterPro; IPR000866; Ahpc-TSA.
 DR InterPro; IPR012336; Thioredoxin-like.
 DR InterPro; IPR012335; Thioredoxin_fold.
 DR Pfam; PF00578; Ahpc-TSA; 1.
 SQ SEQUENCE 195 AA; 21440 MW; 4259767C92A9D31 CRC64;

Query Match 92.1%; Score 943; DB 2; Length 195;
 Best Local Similarity 90.8%; Pred. No. 1,1e-74;
 Matches 177; Conservative 7; Mismatches 11; Indels 0; Gaps 0;

QY 1 MPLLITIDQPPAYOLTLIGDLSKVDKQPGDYFTTTSDEHFGKRVVFFWPKDFTFV 60
 DB 1 MSLLITDQPPAYLTLIGDLSKVDKQPGDYFTTTSDEHFGKRVVFFWPKDFTFV 60
 QY 61 CPTEIAAFSKUNDEFEDRDQIIGVSIIDSEFAHFQWRAQHNDLKTLPFPMLSDIKRELISQ 120
 DB 61 CPTEIAAFGKUNDEFEDRDQIIGVSIIDSEFAHFQWRAQHNDLKTLPFPMLSDIKRELISQ 120
 QY 121 AAGVLNADGVADRTFTVDPNNELQFVSATAGSVGRNDEVLRVLDALQSDCLCACMWRK 180
 DB 121 ATGVLANADGVADRTFTVDPNNELQFVSATAGSVGRNDEVLRVLDALQSDCLCACMWRK 180
 QY 181 GDDPTLDAGELLKASA 195
 DB 181 GDDPTLDAGELLKASA 195

RESULT 4
 057413 MYCAV PRELIMINARY; PRT; 195 AA.

AC 057413; 1-1996 (TRENBLrel. 01, Created)
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
 DT 01-FEB-2005 (TRENBLrel. 29, Last annotation update)
 DE Alkyl hydroperoxidase C (Antigen).
 GN Name=ahpc; Synonyms=Avl-3;
 OS Mycobacterium avium.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium;
 OC NCBI_TaxID=1764;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=95327698; PubMed=7604044;
 RA Sherman D.R., Sabo P.J., Hickey M.J., Arain T.M., Mahairas G.G.,
 RA Yuan Y., Barry C.E. III, Stover C.K.;
 RT "Disparate responses to oxidative stress in saprophytic and pathogenic
 RT Mycobacteria";
 RL Proc. Natl. Acad. Sci. U.S.A. 92:6625-6629(1995).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=92175967; PubMed=1371765;
 RA Yamaguchi R., Matsuo K., Yamazaki A., Takahashi M., Fukasawa Y.,
 RA Wada M., Abe C.;
 RT "Cloning and expression of the gene for the Avl-3 antigen of
 RT Mycobacterium avium and mapping of its epitopes";
 RL Infect. Immun. 60:1210-1216(1992).
 DR EMBL; U18263; AAA79917.1; -; Genomic DNA.
 DR EMBL; M74232; AAA25357.1; -; Genomic DNA.
 DR HSP; P32119; 10MV.
 DR SMK; Q57413; 2-179.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0004601; F:peroxidase activity; IEA.
 DR InterPro; IPR000866; Ahpc-TSA.
 DR InterPro; IPR012336; Thioredoxin-like.
 DR InterPro; IPR012335; Thioredoxin_fold.
 DR Pfam; PF00578; Ahpc-TSA; 1.
 KW Peroxidase.

SQ SEQUENCE 195 AA; 21655 MW; 11A60C1849343FD7 CRC64;

Query Match 91.3%; Score 935; DB 2; Length 195;
 Best Local Similarity 90.3%; Pred. No. 5,7e-74;
 Matches 176; Conservative 6; Mismatches 13; Indels 0; Gaps 0;

QY 1 MPLLITDQPPAYOLTLIGDLSKVDKQPGDYFTTTSDEHFGKRVVFFWPKDFTFV 60
 DB 1 MPLLITDQPPAYELTLIGDLSKVDKQPGDYFTTTSDEHFGKRVVFFWPKDFTFV 60
 QY 61 CPTEIAAFSKUNDEFEDRDQIIGVSIIDSEFAHFQWRAQHNDLKTLPFPMLSDIKRELISQ 120
 DB 61 CPTEIAATFGKUNDEFEDRDQIIGVSIIDSEFAHFQWRAQHNDLKTLPFPMLSDIKRELISQ 120
 QY 121 AAGVLNADGVADRTFTVDPNNELQFVSATAGSVGRNDEVLRVLDALQSDCLCACMWRK 180
 DB 121 ATGVLANADGVADRTFTVDPNNELQFVSATAGSVGRNDEVLRVLDALQSDCLCACMWRK 180
 QY 181 GDDPTLDAGELLKASA 195
 DB 181 GDDPTLNATELLKASA 195

RESULT 5
 073ZL3 MYCPA PRELIMINARY; PRT; 195 AA.

AC 073ZL3;
 DT 05-JUL-2004 (TRENBLrel. 27, Created)
 DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TRENBLrel. 27, Last annotation update)
 DE Ahpc.
 GN Name=ahpc; Ordered locus names=MAP1589c;
 OS Mycobacterium paratuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium;
 OC Mycobacterium avium complex (MAC).
 OC NCBI_TaxID=1770;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA STRAIN=K10;
 RC Li L., Bannantine J., Zhang Q., Amonsin A., Alt D., Kapur V.;
 RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE017232; AAS03906.1; -; Genomic DNA.
 DR SMK; Q73ZL3; 2-179.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0004601; F:peroxidase activity; IEA.
 DR InterPro; IPR000866; Ahpc-TSA.
 DR InterPro; IPR012336; Thioredoxin-like.
 DR Pfam; PF00578; Ahpc-TSA; 1.
 KW Complete proteome.
 SQ SEQUENCE 195 AA; 21641 MW; 4A9D9C0BF102E61D CRC64;

Query Match 91.2%; Score 934; DB 2; Length 195;
 Best Local Similarity 89.7%; Pred. No. 7e-74;
 Matches 175; Conservative 7; Mismatches 13; Indels 0; Gaps 0;

QY 1 MPLLITDQPPAYOLTLIGDLSKVDKQPGDYFTTTSDEHFGKRVVFFWPKDFTFV 60
 DB 1 MPLLITDQPPAYELTLIGDLSKVDKQPGDYFTTTSDEHFGKRVVFFWPKDFTFV 60
 QY 61 CPTEIAAFSKUNDEFEDRDQIIGVSIIDSEFAHFQWRAQHNDLKTLPFPMLSDIKRELISQ 120
 DB 61 CPTEIAATFGKUNDEFEDRDQIIGVSIIDSEFAHFQWRAQHNDLKTLPFPMLSDIKRELISQ 120
 QY 121 AAGVLNADGVADRTFTVDPNNELQFVSATAGSVGRNDEVLRVLDALQSDCLCACMWRK 180
 DB 121 ATGVLANADGVADRTFTVDPNNELQFVSATAGSVGRNDEVLRVLDALQSDCLCACMWRK 180
 QY 181 GDDPTLDAGELLKASA 195
 DB 181 GDDPTLNATELLKASA 195

```
RESULT 6
Q9AGF6_MYCPA PRELIMINARY; PRT; 195 AA.
ID Q9AGF6_MYCPA PRELIMINARY; PRT; 195 AA.
AC Q9AGF6;
DT 01-JUN-2001 (TREMblrel. 17, Created)
DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
DT 01-OCT-2003 (TREMblrel. 25, Last annotation update)
DE Alkylhydroperoxidase C.
GN Name=ahpc;
OS Mycobacterium paratuberculosis.
OC Bacteria; Actinobacteria; Actinomycetales;
OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium avium complex (MAC).
OX NCBI_TaxID=1770;
RN NUCLEOTIDE SEQUENCE.
RP Chang Y.F., Shin K.-S.;
RA Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
RL EMBL; AF34163; AAK20392.1; -; Genomic_DNA.
DR HSSP; Q9T2X2; 1E2Y.
DR SMK; Q9AGF6; 2-179.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0004601; F:peroxidase activity; IEA.
DR InterPro; IPR012336; Thioresdoxin-like.
DR InterPro; IPR012335; Thioresdoxin-fold.
DR Pfam; PF00578; Ahpc-TSA; 1.
DR Peroxidase.
KW SEQUENCE 195 AA; 21726 MW; 808557F1FA5F78B9 CRC64;

Query Match 89.8%; Score 920; DB 2; Length 195;
Best Local Similarity 88.2%; Pred. No. 1,2e-72;
Matches 172; Conservative 8; Mismatches 15; Indels 0; Gaps 0;

QY 1 MLLTTIGDQPPAYQLTALIGDLSKVDKQPGDYFTTTSDEHPGKRVVFPMPKDTFV 60
DB 1 MLLTTIGDQPPAYQLTALIGDLSKVDKQPGDYFTTTSDEHPGKRVVFPMPKDTFV 60
QY 61 CPTETIAAFSKLNDEFEDRDAQIIQVSIIDSEFAHFQWRQAHNDLKTLPFPMLSIKRELST 120
DB 61 CPTETIAAFSKLNDEFEDRDAQIIQVSIIDSEFAHFQWRQAHNDLKTLPFPMLSIKRELST 120
QY 121 AAGVLNADGVADRTFTFVDPNNEIQFVSATAGSVGRNVDEVLTATLALQSDQLCACMWRK 180
DB 121 AAGVLNADGVADRTFTFVDPNNEIQFVSATAGSVGRNVDEVLTATLALQSDQLCACMWRK 180
QY 181 GDPTLADGELLKASA 195
DB 181 GDPTLADGELLKASA 195

RESULT 7
Q5YTS4_NOCFA PRELIMINARY; PRT; 195 AA.
ID Q5YTS4_NOCFA PRELIMINARY; PRT; 195 AA.
AC Q5YTS4;
DT 25-OCT-2004 (TREMblrel. 28, Created)
DT 25-OCT-2004 (TREMblrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMblrel. 28, Last annotation update)
DE Putative alkylhydroperoxide reductase.
GN Name=ahpc; OrderedLocustNames=nf37890;
OS Nocardiopsis farcinica.
OC Bacteria; Actinobacteria; Actinomycetales;
OC Corynebacteriaceae; Nocardiaceae; Nocardiaceae;
OX NCBI_TaxID=37329;
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=IFM 10152;
RA PubMed=15466710; DOI=10.1073/pnas.0406410101;
RA Ishikawa J., Yamashita A., Mikami Y., Hoshino Y., Kurita H., Hotta K.,
RA Shiba T., Hattori M.;
RT "The complete genomic sequence of Nocardiopsis farcinica IFM 10152.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:14925-14930(2004).
DR EMBL; AP006618; BAD58637.1; -; Genomic_DNA.
```

```
DR SMK; Q5YTS4; 3-179.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0004601; F:peroxidase activity; IEA.
DR InterPro; IPR000866; Ahpc-TSA.
DR InterPro; IPR012336; Thioresdoxin-like.
DR Pfam; PF00578; Ahpc-TSA; 1.
KW Complete proteome.
SQ SEQUENCE 195 AA; 21456 MW; 6B3D2F9DC95910C9 CRC64;

Query Match 89.2%; Score 913; DB 2; Length 195;
Best Local Similarity 86.6%; Pred. No. 4.9e-72;
Matches 168; Conservative 12; Mismatches 14; Indels 0; Gaps 0;

QY 1 MLLTTIGDQPPAYQLTALIGDLSKVDKQPGDYFTTTSDEHPGKRVVFPMPKDTFV 60
DB 1 MLLTTIGDQPPAYQLTALIGDLSKVDKQPGDYFTTTSDEHPGKRVVFPMPKDTFV 60
QY 61 CPTETIAAFSKLNDEFEDRDAQIIQVSIIDSEFAHFQWRQAHNDLKTLPFPMLSIKRELST 120
DB 61 CPTETIAAFSKLNDEFEDRDAQIIQVSIIDSEFAHFQWRQAHNDLKTLPFPMLSIKRELST 120
QY 121 AAGVLNADGVADRTFTFVDPNNEIQFVSATAGSVGRNVDEVLTATLALQSDQLCACMWRK 180
DB 121 AAGVLNADGVADRTFTFVDPNNEIQFVSATAGSVGRNVDEVLTATLALQSDQLCACMWRK 180
QY 181 GDPTLADGELLKASA 194
DB 181 GDPTLADGELLKASA 194

RESULT 8
Q9CBF5_MYCLE PRELIMINARY; PRT; 195 AA.
ID Q9CBF5_MYCLE PRELIMINARY; PRT; 195 AA.
AC Q9CBF5;
DT 01-JUN-2001 (TREMblrel. 17, Created)
DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
DT 01-JUN-2003 (TREMblrel. 24, Last annotation update)
DE Alkyl hydroperoxide reductase.
GN Name=ahpc; OrderedLocustNames=ML2042;
OS Mycobacterium leprae.
OC Bacteria; Actinobacteria; Actinomycetales;
OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium;
OX NCBI_TaxID=1769;
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=TN;
RA MEDLINE=2118732; PubMed=11234002; DOI=10.1038/35059006;
RA Cole S.T., Eigemeier K., Parkhill J., James K.D., Thomson N.R.,
RA Wheeler P.R., Honore N., Garnier T., Churcher C.M., Harris D.E.,
RA Mungall K.L., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R.M., Devlin K., Duthoy S., Feltwell T., Frazer A., Hamlin N.,
RA Holroyd S., Hornsby T., Jags K., Lacroix C., Maclean J., Moule S.,
RA Murphy L.D., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
RA Barrett B.G.;
RT "Passive gene decay in the leprosy bacillus.";
RL Nature 409:1007-1011(2001).
DR EMBL; AL583924; CAC30997.1; -; Genomic_DNA.
DR PIR; E87164; E87164.
DR HSSP; Q9T2X2; 1E2Y.
DR SMK; Q9CBF5; 3-179.
DR Leproma; ML2042; -.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0004601; F:peroxidase activity; IEA.
DR InterPro; IPR000866; Ahpc-TSA.
DR InterPro; IPR012336; Thioresdoxin-like.
DR InterPro; IPR012335; Thioresdoxin-fold.
DR Pfam; PF00578; Ahpc-TSA; 1.
KW Complete proteome.
SQ SEQUENCE 195 AA; 21497 MW; DBB60152F813CEA1 CRC64;

Query Match 88.3%; Score 904; DB 2; Length 195;
```

Best Local Similarity 86.7%; Pred. No. 3e-71;
Matches 169; Conservative 12; Mismatches 14; Indels 0; Gaps 0;

Qy 1 MLLTTIGDQPPAYQTLTALIGDLSKVDKAKQDGYFTTTSDEHPGKRVVFVFWPMDPTFV 60
Db 1 MSLTISIGQPPAYQTLTALIGDLSKVDKAKQDGYFTTTSDEHPGKRVVFVFWPMDPTFV 60

Qy 61 CPTETIAAFSKLNDFFERDQAQILGVSIDSEFAHFQWRAQHNDLKTLPFPMLSDIKRELISQ 120
Db 61 CPTETIAAFSKLNDFFERDQAQILGVSIDSEFAHFQWRAQHNDLKTLPFPMLSDIKRELISA 120

Qy 121 AAGVLNADGVADRTFTYVDPNNELQFVSATAGSVGRNVDEVLRVLDALQSDCLCACMWRK 180
Db 121 AAGVLNADGVADRTFTYVDPNNELQFVSATAGSVGRNVDEVLRVLDALQSDCLCACMWRK 180

Qy 181 GDPTLDAGELLKASA 195
Db 181 GDPTLDAGELLKASA 195

RESULT 9

057529 MYCSM PRELIMINARY; PRT; 195 AA.
AC 057529;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-FEB-2005 (TREMblrel. 29, Last annotation update)
DE ALyl1 hydroperoxide reductase C.
OS Name=ahpc; Synonyms=Ahpc;
OC Mycobacterium smegmatis.
OC Plasmid pDP81.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxId=1772;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=mc2 155; PLASMID=pDP81;
RX MEDLINE=96256622; PubMed=8655566;
RA Dhandayuthapani S., Zhang Y., Mudd M.H., Deretic V.;
RT "Oxidative stress response and its role in sensitivity to isoniazid in
RT Mycobacteria: characterization and inducibility of ahpc by peroxides
RT in Mycobacterium smegmatis and lack of expression in M. aurum and M.
RT tuberculosis";
RL J. Bacteriol. 178:3641-3649(1996).
DR EMBL; U43179; AAC44139.1; -; Genomic DNA.
DR HSSP; P32119; 1QMV.
DR HSSP; P32119; AAC44148.1; -; Genomic DNA.
DR HSSP; P32119; 1QMV.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0004601; F:peroxidase activity; IEA.
DR InterPro; IPR000866; Ahpc-TSA.
DR InterPro; IPR012336; Thioresdoxin-like.
DR InterPro; IPR012335; Thioresdoxin_fold.
DR Pfam; PF00578; Ahpc-TSA; 1.
KW Plasmid.
SQ SEQUENCE 195 AA; 21626 MW; 556A5C17A5DA937F CRC64;

Query Match 85.7%; Score 878; DB 2; Length 195;
Best Local Similarity 82.2%; Pred. No. 5.8e-69;
Matches 157; Conservative 20; Mismatches 14; Indels 0; Gaps 0;

Qy 1 MLLTTIGDQPPAYQTLTALIGDLSKVDKAKQDGYFTTTSDEHPGKRVVFVFWPMDPTFV 60
Db 1 MALTLTIGDQPPAYQTLTALIGDLSKVDKAKQDGYFTTTSDEHPGKRVVFVFWPMDPTFV 60

Qy 61 CPTETIAAFSKLNDFFERDQAQILGVSIDSEFAHFQWRAQHNDLKTLPFPMLSDIKRELISQ 120
Db 61 CPTETIAAFSKLNDFFERDQAQILGVSIDSEFAHFQWRAQHNDLKTLPFPMLSDIKRELISA 120

Qy 121 AAGVLNADGVADRTFTYVDPNNELQFVSATAGSVGRNVDEVLRVLDALQSDCLCACMWRK 180
Db 121 AAGVLNADGVADRTFTYVDPNNELQFVSATAGSVGRNVDEVLRVLDALQSDCLCACMWRK 180

Qy 181 GDPTLDAGELL 191
Db 181 GDPTLDAGELL 191

RESULT 10

06NGT3 CORDI PRELIMINARY; PRT; 198 AA.
AC 06NGT3;
DT 05-JUN-2004 (TREMblrel. 27, Created)
DT 05-JUN-2004 (TREMblrel. 27, Last sequence update)
DT 05-JUN-2004 (TREMblrel. 27, Last annotation update)
DE Iron repressible polypeptide (putative reductase).
GN Name=dira; OrderedLocNames=DIP1420;
OS Corynebacterium diptheriae.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxId=1717;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Biotype gravis / NCTC 13129;
RX MEDLINE=22965443; PubMed=14602910; DOI=10.1093/nar/gk9874;
RA Cerdano-Tarraga A.-M., Efstathiou A., Dover L.G., Holden M.T.G.,
RA Pallen M.J., Bentley S.D., Beara G.S., Churcher C.M., James K.D.,
RA De Zeyva N., Chillingworth T., Cronin A., Dowd L., Fellwell T.,
RA Hamlin N., Holroyd S., Jagels K., Moule S., Quail M.A.,
RA Rabinowitsch E., Rutherford K.M., Thomson N.R., Unwin L.,
RA Whitehead S., Barrett B.G., Parkhill J.;
RT "The complete genome sequence and analysis of Corynebacterium
RT diptheriae NCTC13129";
RL Nucleic Acids Res. 31:6516-6523 (2003).
DR EMBL; BX248358; CAE49951.1; -; Genomic DNA.
DR InterPro; IPR000866; Ahpc-TSA.
DR InterPro; IPR012336; Thioresdoxin-like.
DR Pfam; PF00578; Ahpc-TSA; 1.
KW Complete proteome.
SQ SEQUENCE 198 AA; 22365 MW; 7A09C8F1B43205DD CRC64;

Query Match 67.7%; Score 693; DB 2; Length 198;
Best Local Similarity 67.9%; Pred. No. 1.1e-52;
Matches 125; Conservative 29; Mismatches 30; Indels 0; Gaps 0;

Qy 1 MLLTTIGDQPPAYQTLTALIGDLSKVDKAKQDGYFTTTSDEHPGKRVVFVFWPMDPTFV 60
Db 1 MSLTISIGQPPAYQTLTALIGDLSKVDKAKQDGYFTTTSDEHPGKRVVFVFWPMDPTFV 60

Qy 61 CPTETIAAFSKLNDFFERDQAQILGVSIDSEFAHFQWRAQHNDLKTLPFPMLSDIKRELISQ 120
Db 61 CPTETIAAFSKLNDFFERDQAQILGVSIDSEFAHFQWRAQHNDLKTLPFPMLSDIKRELISA 120

Qy 121 AAGVLNADGVADRTFTYVDPNNELQFVSATAGSVGRNVDEVLRVLDALQSDCLCACMWRK 180
Db 121 AAGVLNADGVADRTFTYVDPNNELQFVSATAGSVGRNVDEVLRVLDALQSDCLCACMWRK 180

Qy 181 GDPT 184
Db 181 NDPT 184

RESULT 11

046025 CORDI PRELIMINARY; PRT; 198 AA.
AC 046025;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE Iron repressible polypeptide.
GN Name=dira;
OS Corynebacterium diptheriae.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxId=1717;
RN [1]

```

RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C7;
RX MEDLINE=95286547; PubMed=7768861;
RA Tai S.S., Zhu Y.Y.;
RT "Cloning of a Corynebacterium diphtheriae iron-repressible gene that
RT shares sequence homology with the Ahpc subunit of alkyl hydroperoxide
RT reductase of Salmonella typhimurium.";
RL J. Bacteriol. 177:3512-3517(1995).
DR EMBL; U18620; AAA96946.1; -; Genomic_DNA.
DR HSSP; Q9TX22; 1E2Y.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR InterPro; IPR000866; Ahpc-TSA.
DR InterPro; IPR012336; Thioresdoxin-like.
DR InterPro; IPR012335; Thioresdoxin_fold.
DR Pfam; PF00578; Ahpc-TSA; 1.
SQ SEQUENCE 198 AA; 22312 MW; 7A09DCBA07705DD CRC64;

Query Match 66.9%; Score 685; DB 2; Length 198;
Best Local Similarity 67.4%; Pred. No. 5.3e-52;
Matches 124; Conservative 29; Mismatches 31; Indels 0; Gaps 0;

QY 1 MRLTIGDQFPAVQYLTALIGDLSKVDAKOPGDYFTTTSDEHPKMRVVFMPDPFTVC 60
DB 1 MSILTVGKPEPEFLNTALKGSDLDHVNASQPEDVFETVSLDKYEGKMKVVFPPDPFTVC 60
QY 61 CPTETAAFSKLNDEFEDRDAQILGVSIDSEFAHFQWRAQNDLKTLPFPMLSDIKRELISQA 120
DB 61 CPTETAAFGKLNDEFEDRDAQILGVSIDSEFAHFQWRAQNDLKTLPFPMLSDIKRELISQA 120
QY 121 AAGVLNAGVADRVTFIVDPNNEIOFVSATAGSVGRNDEVLRVLDALQSDLCACMKRK 180
DB 121 ALGVENEGVADRVATFIIDPDGIQFVSVTPDAGVNCNDEVLRVLDALQSEVCACMKQK 180
QY 181 GDPT 184
DB 181 NDPT 184

RESULT 12
ID Q9XSV0_STRVD PRELIMINARY; PRT; 194 AA.
AC Q9XSV0;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 12, Last annotation update)
DE Alkyl hydroperoxide reductase.
GN Name=ahpc;
OS Streptomyces viridosporus.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycinae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxId=67581;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=7A;
RX MEDLINE=20356715; PubMed=10902909;
RA Ramachandran S., Magnuson T.S., Crawford D.L.;
RT "Isolation and analysis of three peroxide sensor regulatory gene
RT homologs ahpc, ahpx and oxvr in Streptomyces viridosporus T7A - a
RT lignocellulose degrading actinomycete.";
RL DNA Seq. 11:51-60(2000).
DR EMBL; AF127576; AAD33340.1; -; Genomic_DNA.
DR HSSP; Q9TX22; 1E2Y.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0004601; F:peroxidase activity; IEA.
DR InterPro; IPR000866; Ahpc-TSA.
DR InterPro; IPR012336; Thioresdoxin-like.
DR InterPro; IPR012335; Thioresdoxin_fold.
DR Pfam; PF00578; Ahpc-TSA; 1.
SQ SEQUENCE 194 AA; 21700 MW; F324A03842BC7AD1 CRC64;

Query Match 60.4%; Score 619; DB 2; Length 194;
Best Local Similarity 62.1%; Pred. No. 3.2e-46;

```

```

Matches 118; Conservative 18; Mismatches 46; Indels 8; Gaps 1;

QY 2 PLTTIGDQFPAVQYLTALIGDLSKVDAKOPGDYFTTTSDEHPKMRVVFMPDPFTVC 61
DB 10 PVLTVGDKPEPEFLTACV-----SLKKESETINHKTYBEKMKIVFAMPDPFTVC 61
QY 62 PTEIAAFSKLNDEFEDRDAQILGVSIDSEFAHFQWRAQNDLKTLPFPMLSDIKRELISQA 121
DB 62 PTEIAAFGKLNDEFEDRDAQILGVSIDSEFAHFQWRAQNDLKTLPFPMMSDIGHELMRA 121
QY 122 AAGVLNAGVADRVTFIVDPNNEIOFVSATAGSVGRNDEVLRVLDALQSDLCACMKRK 181
DB 122 LGIEGDEGFARAVFIVDPNNEIOFVMTAGSVGRNDEVLRVLDALQTBELCFCKMTKG 181
QY 182 DPTDAGELL 191
DB 182 DETIDPVAL 191

RESULT 13
ID Q9FBP5_STRCO PRELIMINARY; PRT; 184 AA.
AC Q9FBP5;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Alkyl hydroperoxide reductase.
GN OrderedLocustNames=SC05032; ORFNames=SCK7.05c;
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycinae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxId=1902;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953; DOI=10.1038/417141a;
RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.B., Quail M.A., Kleser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornby T., Howarth S.,
RA Huang C.-H., Kleser T., Larke L., Murphy L.D., Oliver K., O'Neill S.,
RA Rabinowitch E., Rajandream M.A., Rutherford K.M., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J.R., Bartell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2) ";
RL Nature 417:141-147(2002).
DR EMBL; AL391122; CAC05877.1; -; Genomic_DNA.
DR HSSP; P32119; 1QMV.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0004601; F:peroxidase activity; IEA.
DR InterPro; IPR000866; Ahpc-TSA.
DR InterPro; IPR012336; Thioresdoxin-like.
DR InterPro; IPR012335; Thioresdoxin_fold.
DR Pfam; PF00578; Ahpc-TSA; 1.
KM Complete proteome.
SQ SEQUENCE 184 AA; 20726 MW; F77E2820DA85E341 CRC64;

Query Match 57.9%; Score 593; DB 2; Length 184;
Best Local Similarity 61.4%; Pred. No. 5.3e-44;
Matches 116; Conservative 14; Mismatches 51; Indels 8; Gaps 1;

QY 3 LRLTIGDQFPAVQYLTALIGDLSKVDAKOPGDYFTTTSDEHPKMRVVFMPDPFTVC 62
DB 1 MLTVGDKPEPEFLTACV-----SLKKESETINHKTYBEKMKIVFAMPDPFTVC 52
QY 63 TEIAAFSKLNDEFEDRDAQILGVSIDSEFAHFQWRAQNDLKTLPFPMLSDIKRELISQA 122
DB 53 TEIAAFGKLNDEFEDRDAQILGVSIDSEFAHFQWRAQNDLKTLPFPMMSDIGHELMRL 112
QY 123 GVLNAGVADRVTFIVDPNNEIOFVSATAGSVGRNDEVLRVLDALQSDLCACMKRK 182
DB 123 GVLNAGVADRVTFIVDPNNEIOFVSATAGSVGRNDEVLRVLDALQSDLCACMKRK 182

```

```
Db 113 GIEGEDGFAQRAVFIVDONNEIQFTMTAGSVGNPKSVLRAVLDALOTDELCPCKWMSKGD 172
Qy 183 PTLDAGELL 191
Db 173 ETLDPVALL 181

RESULT 14
Q9RNT2_STRCO PRELIMINARY; PRT; 184 AA.
ID Q9RNT2_STRCO
AC Q9RNT2_STRCO
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE Ahpc.
GN Name=ahpc;
OS Streptomyces coelicolor A3(2).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycinae; Streptomycetaceae; Streptomyces.
CX NCBI_TaxID=100226;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=A3;
RX MEDLINE=22206464; Pubmed=12218006;
RX DOI=10.1128/JB.184.19.5214-5222.2002;
RA Hahn J.S., Oh S.Y., Roe J.H.;
RT "Role of Oxyr as a peroxide-sensing positive regulator in Streptomyces
RT coelicolor A3(2).";
RT J. Bacteriol. 184:5214-5222(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=A3;
RA Hahn J.-S., Roe J.-H.;
RL Submitted (SEP-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF186371; AAF06744.1; -; Genomic_DNA.
DR HSSP; P32119; IQNV.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0004601; P:peroxidase activity; IEA.
DR InterPro; IPR000866; Ahpc-TSA.
DR InterPro; IPR012335; Thioresdoxin-like.
DR InterPro; IPR012335; Thioresdoxin_fold.
DR Pfam; PF00578; Ahpc-TSA; 1.
SQ SEQUENCE 184 AA; 20712 MW; E291762EC035E2EA CRC64;

Query Match 57.8%; Score 592; DB 2; Length 184;
Best Local Similarity 61.4%; Pred. No. 7.2e-44;
Matches 116; Conservative 14; Mismatches 51; Indels 8; Gaps 1;

Qy 3 LTTIGDQFPAYQTLALIGDLSKYDAKQPGDYFTTITSDEHPGKRVVFFMPKQFTFVCP 62
Db 1 MLTVGDKRPFEPDLTACV-----SLEKGRDFOQINHKTVEGQKVFAMPKQFTFVCP 52

Qy 63 TEIAAFKLNDEFEDRDQAQIIIGVSIIDSEFAHFQWRQAQNDLKTLPFPMLSDIKELSOAA 122
Db 53 TEIAAFKLNDEFEDRDQAQIIIGVSGDSEFVHNAWKCHDRLDLPFPMADSKEHLMRDL 112

Qy 123 GVINADGVADRVTFIVDPNNEIQFVSATAGSVGNRVNDEVLRVLDALOSDELCAQWNRKD 182
Db 113 GIEGEDGFAQRAVFIVDONNEIQFTMTAGSVGNPKSVLRAVLDALOTDELCPCKWMSKGD 172

Qy 183 PTLDAGELL 191
Db 173 ETLDPVALL 181

RESULT 15
Q82IC6_STRAM PRELIMINARY; PRT; 184 AA.
ID Q82IC6_STRAM
AC Q82IC6_STRAM
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Putative alkyl hydroperoxide reductase.
```

```
GN Name=ahpc; OrderedLocusNames=SAV3232;
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycinae; Streptomycetaceae; Streptomyces.
CX NCBI_TaxID=33903;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=22608306; Pubmed=12692562; DOI=10.1038/nbt820;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T., Sakaki Y., Hattori M., Omura S.;
RT "Complete genome sequence and comparative analysis of the industrial microorganism Streptomyces avermitilis.";
RL Nat. Biotechnol. 21:526-531(2003).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=21477403; Pubmed=11572948; DOI=10.1073/pnas.21133198;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C., Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osone T., Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT "Genome sequence of an industrial microorganism Streptomyces avermitilis: deducing the ability of producing secondary metabolites.";
RT Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
DR EMBL; BA000030; BAC70943.1; -; Genomic_DNA.
DR HSSP; Q9TXZ2; IE2Y.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0004601; P:peroxidase activity; IEA.
DR InterPro; IPR000866; Ahpc-TSA.
DR InterPro; IPR012336; Thioresdoxin-like.
DR InterPro; IPR012335; Thioresdoxin_fold.
DR Pfam; PF00578; Ahpc-TSA; 1.
KW Complete proteome.
SQ SEQUENCE 184 AA; 20773 MW; 6F98071259D81088 CRC64;

Query Match 57.6%; Score 590; DB 2; Length 184;
Best Local Similarity 59.3%; Pred. No. 1.1e-43;
Matches 112; Conservative 21; Mismatches 48; Indels 8; Gaps 1;

Qy 3 LTTIGDQFPAYQTLALIGDLSKYDAKQPGDYFTTITSDEHPGKRVVFFMPKQFTFVCP 62
Db 1 MLTVGDKRPFEPDLTACV-----SLEKGRDFOQINHKTVEGQKVFAMPKQFTFVCP 52

Qy 63 TEIAAFKLNDEFEDRDQAQIIIGVSIIDSEFAHFQWRQAQNDLKTLPFPMLSDIKELSOAA 122
Db 53 TEIAAFKLNDEFEDRDQAQIIIGVSGDSEFVHNAWKCHDRLDLPFPMADSKEHLMRDL 112

Qy 123 GVINADGVADRVTFIVDPNNEIQFVSATAGSVGNRVNDEVLRVLDALOSDELCAQWNRKD 182
Db 113 GIEGEDGFAQRAVFIVDONNEIQFVSATAGSVGNPKSVLRAVLDALOTDELCPCKWMSKGD 172

Qy 183 PTLDAGELL 191
Db 173 ETLDPVALL 181

Search completed: March 23, 2006, 05:24:25
Job time : 50.8737 secs
```

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OW protein - protein search, using sw model

Run on: March 23, 2006, 05:24:53 ; Search time 13.2997 Seconds
(without alignments)
1212.194 Million cell updates/sec

Title: US-10-617-038-26
Perfect score: 1024
Sequence: 1 MPLLTIGDPFAYQLTALIG.....CNMRKDPDLDELKASA 195

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents, AA:
1: /cgn2_6/ptodata/1/1aa/5_COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/6_COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/H_COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/PCITIS_COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/RE_COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	898	87.7	195	2	US-08-311-731A-84 Sequence 84, Appl
2	524.5	51.2	184	2	US-09-902-540-10665 Sequence 10665, A
3	350.5	34.2	982	2	US-09-551-974A-95 Sequence 95, Appl
4	350.5	34.2	982	2	US-09-565-501A-95 Sequence 95, Appl
5	350.5	34.2	982	2	US-09-639-206A-95 Sequence 95, Appl
6	350.5	34.2	982	2	US-09-874-923-95 Sequence 95, Appl
7	350.5	34.2	1427	2	US-09-551-974A-97 Sequence 97, Appl
8	350.5	34.2	1427	2	US-09-565-501A-97 Sequence 97, Appl
9	350.5	34.2	1427	2	US-09-639-206A-97 Sequence 97, Appl
10	350.5	34.2	1427	2	US-09-874-923-97 Sequence 97, Appl
11	350.5	34.2	1464	2	US-10-012-886-1008 Sequence 1008, Ap
12	350.5	34.2	1641	2	US-09-551-974A-96 Sequence 96, Appl
13	350.5	34.2	1641	2	US-09-565-501A-96 Sequence 96, Appl
14	350.5	34.2	1641	2	US-09-639-206A-96 Sequence 96, Appl
15	350.5	34.2	1641	2	US-09-874-923-96 Sequence 96, Appl
16	346.5	33.8	199	2	US-09-183-861-24 Sequence 24, Appl
17	346.5	33.8	199	2	US-09-022-765-24 Sequence 24, Appl
18	346.5	33.8	199	2	US-09-551-974A-24 Sequence 24, Appl
19	346.5	33.8	199	2	US-09-565-501A-24 Sequence 24, Appl
20	346.5	33.8	199	2	US-09-639-206A-24 Sequence 24, Appl
21	346.5	33.8	199	2	US-08-874-923-24 Sequence 24, Appl
22	346.5	33.8	199	2	US-08-798-841-24 Sequence 24, Appl
23	344	33.6	206	2	US-09-183-861-26 Sequence 26, Appl
24	344	33.6	206	2	US-09-022-765-26 Sequence 26, Appl
25	344	33.6	206	2	US-09-551-974A-26 Sequence 26, Appl
26	344	33.6	206	2	US-09-565-501A-26 Sequence 26, Appl
27	344	33.6	206	2	US-09-639-206A-26 Sequence 26, Appl

28	344	33.6	206	2	US-09-874-923-26 Sequence 26, Appl
29	344	33.6	206	2	US-08-798-841-26 Sequence 26, Appl
30	327.5	32.0	244	2	US-09-252-991A-29692 Sequence 29692, A
31	323.5	31.6	205	2	US-09-489-039A-13084 Sequence 13084, A
32	320.5	31.3	199	1	US-08-467-265-17 Sequence 17, Appl
33	320.5	31.3	199	1	US-08-467-265-17 Sequence 17, Appl
34	320.5	31.3	199	2	US-09-407-891-17 Sequence 1055, Ap
35	318	31.1	206	2	US-09-538-092-1055 Sequence 8377, Ap
36	318	31.1	257	1	US-08-467-265-16 Sequence 16, Appl
37	318	31.1	257	2	US-08-467-265-16 Sequence 16, Appl
38	318	31.1	257	2	US-09-407-891-16 Sequence 16, Appl
39	318	31.1	257	2	US-09-375-907-5 Sequence 5, Appl1
40	318	31.1	257	2	US-09-270-767-46784 Sequence 1039, Ap
41	317	31.0	248	2	US-09-538-092-1039 Sequence 6760, Ap
42	317	31.0	256	2	US-09-949-016-6760 Sequence 10249, A
43	317	31.0	256	2	US-09-949-016-6760 Sequence 10249, A
44	317	31.0	258	2	US-09-949-016-10249 Sequence 4, Appl1
45	315.5	30.8	199	2	US-09-375-907-4 Sequence 4, Appl1

ALIGNMENTS

RESULT 1
US-08-311-731A-84
; Sequence 84, Application US/08311731A
; Patent No. 6583266
; GENERAL INFORMATION:
; APPLICANT: SMITH, DOUGLAS
; APPLICANT: MAO, JEN-I
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; TITLE OF INVENTION: RELATING TO MYCOBACTERIUM TUBERCULOSIS AND LAPRAE FOR
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 411
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: WOLF, GREENFIELD & SACKS, P. C.
; STREET: 600 ATLANTIC AVENUE
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 02210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/311,731A
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: GATES, EDWARD R.
; REGISTRATION NUMBER: 31,616
; REFERENCE/DOCKET NUMBER: C0044/7125
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/720-3500
; TELEFAX: 617/720-2441
; INFORMATION FOR SEQ ID NO: 84:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 195 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: MYCOBACTERIUM LAPRAE
; US-08-311-731A-84
Query Match 87.7%; Score 898; DB 2; Length 195;
Best Local Similarity 85.6%; Pred. No. 1.6e-100;
Matches 167; Conservative 14; Mismatches 0; Gaps 0;
QY 1 MPLLTIGDPFAYQLTALIGDLSKVDKAGDYFTTTSDEHFGKMRVVFPMKDPFTFV 60

```
Db 1 MSLSIGQFPAYQTLALIGDLSKVDKQPGDYFTTISDEHPGKRWVFFPMPDFTFPI 60
Qy 61 CPTETAAASKLNDSEFEDRDAQILGVSIDSEFAHPQWRAQHNDLKTLPPEMLSDIKRELSQL 120
Db 61 CPTETAAASKLNDSEFEDRDAQILGVSIDSEFAHPQWRAQHNDLKTLPPEMLSDIKRELSQL 120
Qy 121 AAGVADGVDRTFTIYDPNNEIQFVSATAGSVGRNDEVLRVLDALQSDCLCACNMRK 180
Db 121 AAGVADGVDRTFTIYDPNNEIQFVSATAGSVGRNDEVLRVLDALQSDCLCACNMRK 180
Qy 181 GDPTLDAGELLKASA 195
Db 181 GDPTLDAGELLKASA 195
```

RESULT 2

```
US-09-902-540-10665
; Sequence 10665, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 10665
; LENGTH: 184
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-10665
```

```
Query Match 51.2%; Score 524.5; DB 2; Length 184;
Best Local Similarity 54.3%; Pred. No. 3e-55;
Matches 100; Conservative 25; Mismatches 50; Indels 9; Gaps 2;
```

```
Qy 3 LITIGDQFPAYQTLALIGDLSKVDKQPGDYFTTISDEHPGKRWVFFPMPDFTFVCP 62
Db 1 MLTIGDQFPAYQTLALIGDLSKVDKQPGDYFTTISDEHPGKRWVFFPMPDFTFVCP 62
Qy 63 TEIAFSKLNDEFEEDRDAQILGVSIDSEFAHPQWRAQHNDLKTLPPEMLSDIKRELSQL 122
Db 53 TEIAFSKLNDEFEEDRDAQILGVSIDSEFAHPQWRAQHNDLKTLPPEMLSDIKRELSQL 122
Qy 123 GVLN-ADGVADRTFTIYDPNNEIQFVSATAGSVGRNDEVLRVLDALQSDCLCACNMRK 181
Db 113 GILHKEBVALRAFTIYDPNNEIQFVSATAGSVGRNDEVLRVLDALQSDCLCACNMRK 172
Qy 182 DPTL 185
Db 173 ERTL 176
```

RESULT 3

```
US-09-551-974A-95
; Sequence 95, Application US/09551974A
; Patent No. 6500437
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Webb, John R.
; APPLICANT: Dillon, David C.
; APPLICANT: Skeiky, Yasir A.W.
; TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USE IN THE
; FILE REFERENCE: 210121.420C5
; CURRENT APPLICATION NUMBER: US/09/551.974A
```

```
; CURRENT FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 101
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 95
; LENGTH: 982
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion (poly-protein) constructs comprising multiple
US-09-551-974A-95
```

```
Query Match 34.2%; Score 350.5; DB 2; Length 982;
Best Local Similarity 39.4%; Pred. No. 5.2e-33;
Matches 74; Conservative 35; Mismatches 60; Indels 19; Gaps 4;
```

```
Qy 6 IGDFPAYQTLALIGDLSKVDKQPGDYFTTISDEHPGKRWVFFPMPDFTFVCPTEI 65
Db 15 INSPAPSEFVALM-----PNSFKKISLSYKGMVVLFFYPIDFTFVCPTEV 63
Qy 66 AAFSKLNDEFEEDRDAQILGVSIDSEFAHPQWRAQHNDLKTLPPEMLSDIKRELSQL 122
Db 64 IAFSDSVSRFNLNCEVLACISDEVAHLQMTLQDRKKGLGTMALPMLADTKTSIARSY 123
Qy 123 GVL-NADGVADRTFTIYDPNNEIQFVSATAGSVGRNDEVLRVLDALQ-----SDELCA 177
Db 124 GVLEESQGVAVYRGIFIDPHGMRLQITVNDMPGVRSVEVLRLEAFQVFKHGEVCPAN 183
Qy 178 WRKGDPTL 185
Db 184 WKKGAPTM 191
```

RESULT 4

```
US-09-565-501A-95
; Sequence 95, Application US/09565501A
; Patent No. 6607731
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Webb, John R.
; APPLICANT: Dillon, David C.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Bhatia, Ajay
; APPLICANT: Coler, Rhea
; APPLICANT: Peter Probst
; TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USE IN THE
; FILE REFERENCE: 210121.420C6
; CURRENT APPLICATION NUMBER: US/09/565.501A
; CURRENT FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 95
; LENGTH: 982
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion (poly-protein) constructs comprising multiple
US-09-565-501A-95
```

```
Query Match 34.2%; Score 350.5; DB 2; Length 982;
Best Local Similarity 39.4%; Pred. No. 5.2e-33;
Matches 74; Conservative 35; Mismatches 60; Indels 19; Gaps 4;
```

```
Qy 6 IGDFPAYQTLALIGDLSKVDKQPGDYFTTISDEHPGKRWVFFPMPDFTFVCPTEI 65
Db 15 INSPAPSEFVALM-----PNSFKKISLSYKGMVVLFFYPIDFTFVCPTEV 63
Qy 66 AAFSKLNDEFEEDRDAQILGVSIDSEFAHPQWRAQHNDLKTLPPEMLSDIKRELSQL 122
Db 64 IAFSDSVSRFNLNCEVLACISDEVAHLQMTLQDRKKGLGTMALPMLADTKTSIARSY 123
```



```
Db      64 IAFSDSVSRFNEELNCEVLACSIDSEYAHLOWTLOQRKKGGIGTMAIPMLADTKSIASYS 123
Qy      123 GVL-NADGVADRVTFIYDPNNEIOFVSATAGSVGRNDEVLRVLDALQ----SDELCACN 177
Db      124 GVLSESOQVAVRGLFIIDPHGMLRQITVNDMPVGRSVBEVRLLEAFQFVEKGEVCPAN 183
Qy      178 WRKGDPTL 185
Db      184 WKKGAPTM 191

RESULT 8
US-09-565-501A-97
; Sequence 97, Application US/09565501A
; Patent No. 6607731
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Webb, John R.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Bhatia, Ajay
; APPLICANT: Coler, Rhea
; APPLICANT: Peter Probst
; TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USE IN THE
; TITLE OF INVENTION: THERAPY AND DIAGNOSIS OF LEISHMANIASIS
; FILE REFERENCE: 210121.420C6
; CURRENT APPLICATION NUMBER: US/09/565,501A
; CURRENT FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 97
; LENGTH: 1427
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion (poly-protein) constructs comprising multiple
; OTHER INFORMATION: Leishmania antigens
US-09-565-501A-97

Query Match      34.2%; Score 350.5; DB 2; Length 1427;
Best Local Similarity 39.4%; Pred. No. 9.3e-33;
Matches 74; Conservative 35; Mismatches 60; Indels 19; Gaps 4;

Qy      6 IGDOPAYQLTALIGDLSKVDAKQPGDYFTTITSDSHPGKRVVFPMKDFVFCPTREI 65
Db      15 INSPAPSFEEVYALM-----PNSFKKISLSYKGMVVLFFYPIDFTFVCPTREV 63
Qy      66 AAFSKLDEFEDRDQAQILGVSIDSEFAHFQWRAQ---HNDLKTLPFPMLSDIKRELQAA 122
Db      64 IAFSDSVSRFNEELNCEVLACSIDSEYAHLOWTLOQRKKGGIGTMAIPMLADTKSIASYS 123
Qy      123 GVL-NADGVADRVTFIYDPNNEIOFVSATAGSVGRNDEVLRVLDALQ----SDELCACN 177
Db      124 GVLSESOQVAVRGLFIIDPHGMLRQITVNDMPVGRSVBEVRLLEAFQFVEKGEVCPAN 183
Qy      178 WRKGDPTL 185
Db      184 WKKGAPTM 191

RESULT 9
US-09-639-206A-97
; Sequence 97, Application US/09639206A
; Patent No. 6613337
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Webb, John R.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Bhatia, Ajay
```

```
; APPLICANT: Coler, Rhea
; APPLICANT: Probst, Peter
; TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USE IN THE
; TITLE OF INVENTION: THERAPY AND DIAGNOSIS OF LEISHMANIASIS
; FILE REFERENCE: 210121.420C7
; CURRENT APPLICATION NUMBER: US/09/639,206A
; CURRENT FILING DATE: 2000-08-14
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 97
; LENGTH: 1427
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion (poly-protein) constructs comprising multiple
; OTHER INFORMATION: Leishmania antigens
US-09-639-206A-97

Query Match      34.2%; Score 350.5; DB 2; Length 1427;
Best Local Similarity 39.4%; Pred. No. 9.3e-33;
Matches 74; Conservative 35; Mismatches 60; Indels 19; Gaps 4;

Qy      6 IGDOPAYQLTALIGDLSKVDAKQPGDYFTTITSDSHPGKRVVFPMKDFVFCPTREI 65
Db      15 INSPAPSFEEVYALM-----PNSFKKISLSYKGMVVLFFYPIDFTFVCPTREV 63
Qy      66 AAFSKLDEFEDRDQAQILGVSIDSEFAHFQWRAQ---HNDLKTLPFPMLSDIKRELQAA 122
Db      64 IAFSDSVSRFNEELNCEVLACSIDSEYAHLOWTLOQRKKGGIGTMAIPMLADTKSIASYS 123
Qy      123 GVL-NADGVADRVTFIYDPNNEIOFVSATAGSVGRNDEVLRVLDALQ----SDELCACN 177
Db      124 GVLSESOQVAVRGLFIIDPHGMLRQITVNDMPVGRSVBEVRLLEAFQFVEKGEVCPAN 183
Qy      178 WRKGDPTL 185
Db      184 WKKGAPTM 191

RESULT 10
US-09-874-923-97
; Sequence 97, Application US/09874923
; Patent No. 6638517
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Webb, John R.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Bhatia, Ajay
; APPLICANT: Coler, Rhea
; APPLICANT: Probst, Peter
; APPLICANT: Brannon, Mark
; TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USE IN THE
; TITLE OF INVENTION: THERAPY AND DIAGNOSIS OF LEISHMANIASIS
; FILE REFERENCE: 210121.420C8
; CURRENT APPLICATION NUMBER: US/09/874,923
; CURRENT FILING DATE: 2001-06-04
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 97
; LENGTH: 1427
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion (poly-protein) constructs comprising multiple
; OTHER INFORMATION: Leishmania antigens
US-09-874-923-97

Query Match      34.2%; Score 350.5; DB 2; Length 1427;
Best Local Similarity 39.4%; Pred. No. 9.3e-33;
Matches 74; Conservative 35; Mismatches 60; Indels 19; Gaps 4;
```

```
Qy 6 IGDOPAYQLTALIGDLSKVDAKQPGDYFTTITSDEHFGKRRVVFPMKQFTFVCPTREI 65
Db 15 INSPAPSEFEVALM-----PNSGFKKISLSYKGMVVFYPLDFTFVCPTREV 63
Qy 66 AAFSLKDNDEPRDQAIIIGVSIIDSEFAHFQWRAQ--HNDKTLPEFPMLSIKRELSDQA 122
Db 64 IAFSDSVSRFNLNCEVLACSIDSSEYAHLOWTLODRKKGIGTMAIPMLADTKXSIAASY 123
Qy 123 GVL-NADGVADRVTFIVDPNNEIOFVSATAGSVGRNDEVLRVLDALQ----SDELCAQN 177
Db 124 GVLSESGVAYRGLFTIDPHGLRQITVNDMPVGRSVBEVLRLLEAFQFVEKHGVCPCAN 183
Qy 178 WRKGDPTL 185
Db 184 WKKGAPTM 191

RESULT 11
US-10-012-896-1008
; Sequence 1008, Application US/10012896
; Patent No. 6943236
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuguu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darlick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepner, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Basols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; APPLICANT: Wantanabe, Yoshihiro
; APPLICANT: Meagher, Madeleine Joy
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C27
; CURRENT APPLICATION NUMBER: US/10/012,896
; NUMBER OF SEQ ID NOS: 1011
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1008
; LENGTH: 1464
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-012-896-1008

Query Match 34.2%; Score 350.5; DB 2; Length 1464;
Best Local Similarity 39.4%; Pred. No. 9.7e-33;
Matches 74; Conservative 35; Mismatches 60; Indels 19; Gaps 4;
```

```
Db 117 GVLSESGVAYRGLFTIDPHGLRQITVNDMPVGRSVBEVLRLLEAFQFVEKHGVCPCAN 176
Qy 178 WRKGDPTL 185
Db 177 WKKGAPTM 184

RESULT 12
US-09-551-974A-96
; Sequence 96, Application US/09551974A
; Patent No. 6500437
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Webb, John R.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Skeiky, Yasir A.W.
; TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USE IN THE
; FILE REFERENCE: 210121.420C5
; CURRENT APPLICATION NUMBER: US/09/551,974A
; NUMBER OF SEQ ID NOS: 101
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 96
; LENGTH: 1641
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion (poly-protein) constructs comprising multiple
US-09-551-974A-96

Query Match 34.2%; Score 350.5; DB 2; Length 1641;
Best Local Similarity 39.4%; Pred. No. 1.2e-32;
Matches 74; Conservative 35; Mismatches 60; Indels 19; Gaps 4;
```

```

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 96
; LENGTH: 1641
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion (poly-protein) constructs comprising multiple
; OTHER INFORMATION: Leishmania antigens
US-09-565-501A-96

Query Match      34.2%; Score 350.5; DB 2; Length 1641;
Best Local Similarity 39.4%; Pred. No. 1.2e-32;
Matches 74; Conservative 35; Mismatches 60; Indels 19; Gaps 4;

QY 6 IGDQPAYQLTALIGDLSKVDKQPGDYFTTITSDEHKGKRVVFFWPKDFTVCPTPEI 65
DB 15 INSPAPSEFEVYALM-----PNSGFKKISLSYKGMVVLFFYPPLDFTFCPTREV 63
QY 66 AAFSKLNDDEFDRDAQILGVSIDSEFAHFOWRQAQ---HNDLKTLPFPMLSDIKRELSQA 122
DB 64 IAFSDSVSRFVELNCEVLACSIDSEYAHLOWTLDQRKKGGIGTWAIPMLADTKTSIARSY 123
QY 123 GVL-NADGVADRVTFTYDPNNEIOFVSATAGSVGRNDEVLRVLDALQ---SDELCAQN 177
DB 124 GVLSESGVAYRGLFIIDPHGMLRQITVNDMPVGRSVBEVLRLLEAFQFVEKHGEVCPAN 183
QY 178 WRKGDPPTL 185
DB 184 WKKGAPTM 191

RESULT 14
US-09-639-206A-96
; Sequence 96, Application US/09639206A
; Patent No. 6613337
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Campos-Neco, Antonio
; APPLICANT: Webb, John R.
; APPLICANT: Dillon, David C.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Bharti, Ajay
; APPLICANT: Colet, Rhea
; APPLICANT: Probst, Peter
; TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USE IN THE
; FILE REFERENCE: 210121.420C7
; CURRENT APPLICATION NUMBER: US/09/639,206A
; CURRENT FILING DATE: 2000-08-14
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 96
; LENGTH: 1641
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion (poly-protein) constructs comprising multiple
; OTHER INFORMATION: Leishmania antigens
US-09-639-206A-96

Query Match      34.2%; Score 350.5; DB 2; Length 1641;
Best Local Similarity 39.4%; Pred. No. 1.2e-32;
Matches 74; Conservative 35; Mismatches 60; Indels 19; Gaps 4;

QY 6 IGDQPAYQLTALIGDLSKVDKQPGDYFTTITSDEHKGKRVVFFWPKDFTVCPTPEI 65
DB 15 INSPAPSEFEVYALM-----PNSGFKKISLSYKGMVVLFFYPPLDFTFCPTREV 63
QY 66 AAFSKLNDDEFDRDAQILGVSIDSEFAHFOWRQAQ---HNDLKTLPFPMLSDIKRELSQA 122
DB 64 IAFSDSVSRFVELNCEVLACSIDSEYAHLOWTLDQRKKGGIGTWAIPMLADTKTSIARSY 123
QY 123 GVL-NADGVADRVTFTYDPNNEIOFVSATAGSVGRNDEVLRVLDALQ---SDELCAQN 177
DB 124 GVLSESGVAYRGLFIIDPHGMLRQITVNDMPVGRSVBEVLRLLEAFQFVEKHGEVCPAN 183
QY 178 WRKGDPPTL 185
DB 184 WKKGAPTM 191
```

```

DB 124 GVLSESGVAYRGLFIIDPHGMLRQITVNDMPVGRSVBEVLRLLEAFQFVEKHGEVCPAN 183
QY 178 WRKGDPPTL 185
DB 184 WKKGAPTM 191

RESULT 15
US-09-874-923-96
; Sequence 96, Application US/09874923
; Patent No. 6638517
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Campos-Neco, Antonio
; APPLICANT: Webb, John R.
; APPLICANT: Dillon, David C.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Bharti, Ajay
; APPLICANT: Colet, Rhea
; APPLICANT: Probst, Peter
; APPLICANT: Brannon, Mark
; TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USE IN THE
; FILE REFERENCE: 210121.420C8
; CURRENT APPLICATION NUMBER: US/09/874,923
; CURRENT FILING DATE: 2001-06-04
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 96
; LENGTH: 1641
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion (poly-protein) constructs comprising multiple
; OTHER INFORMATION: Leishmania antigens
US-09-874-923-96

Query Match      34.2%; Score 350.5; DB 2; Length 1641;
Best Local Similarity 39.4%; Pred. No. 1.2e-32;
Matches 74; Conservative 35; Mismatches 60; Indels 19; Gaps 4;

QY 6 IGDQPAYQLTALIGDLSKVDKQPGDYFTTITSDEHKGKRVVFFWPKDFTVCPTPEI 65
DB 15 INSPAPSEFEVYALM-----PNSGFKKISLSYKGMVVLFFYPPLDFTFCPTREV 63
QY 66 AAFSKLNDDEFDRDAQILGVSIDSEFAHFOWRQAQ---HNDLKTLPFPMLSDIKRELSQA 122
DB 64 IAFSDSVSRFVELNCEVLACSIDSEYAHLOWTLDQRKKGGIGTWAIPMLADTKTSIARSY 123
QY 123 GVL-NADGVADRVTFTYDPNNEIOFVSATAGSVGRNDEVLRVLDALQ---SDELCAQN 177
DB 124 GVLSESGVAYRGLFIIDPHGMLRQITVNDMPVGRSVBEVLRLLEAFQFVEKHGEVCPAN 183
QY 178 WRKGDPPTL 185
DB 184 WKKGAPTM 191

Search completed: March 23, 2006, 05:30:15
Job time : 13.3997 secs
```

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 23, 2006, 06:14:33 ; Search time 41.0073 Seconds
(without alignments)
1986.885 Million cell updates/sec

Title: US-10-617-038-26

Perfect score: 1024
Sequence: 1 MFLTRIGDQPAYQUTALIG.....CNMRKDPPLDAGELLKASA 195

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA_Main:*
1: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep:*
3: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*
4: /cgn2_6/ptodata/1/pubppaa/US10a_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubppaa/US10b_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubppaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1024	100.0	195	4	US-10-282-122A-62742
2	1024	100.0	195	4	US-10-282-122A-64708
3	1024	100.0	195	4	US-10-617-038-26
4	1024	100.0	195	4	US-10-332-512A-21
5	1024	100.0	195	5	US-10-732-923-21281
6	943	92.1	195	5	US-10-732-923-21020
7	935	91.3	195	4	US-10-282-122A-61886
8	935	91.3	195	5	US-10-732-923-21016
9	920	89.8	195	5	US-10-732-923-21018
10	904	88.3	195	5	US-10-282-122A-63991
11	904	88.3	195	5	US-10-732-923-21017
12	878	85.7	195	5	US-10-732-923-21019
13	693	67.7	198	4	US-10-282-122A-53752
14	685	66.9	198	5	US-10-732-923-21013
15	619	60.4	194	5	US-10-732-923-21258
16	593	57.9	184	5	US-10-732-923-20950
17	592	57.8	184	5	US-10-732-923-20951
18	590	57.6	184	5	US-10-732-923-21095
19	589	57.5	184	4	US-10-156-761-10769
20	569.5	55.6	179	5	US-10-732-923-21094
21	530.5	51.8	188	4	US-10-369-493-19388
22	475	46.4	185	5	US-10-732-923-21306
23	453	43.2	184	5	US-10-732-923-21050
24	444	43.4	184	5	US-10-732-923-21087
25	444	43.4	184	5	US-10-732-923-21088
26	432.5	42.2	182	4	US-10-282-122A-49002
27	432.5	42.2	182	5	US-10-732-923-21322

28	429.5	41.9	230	5	US-10-732-923-21109	Sequence 21109, A
29	420.5	41.1	182	4	US-10-732-923-20992	Sequence 20992, A
30	410.5	40.1	182	4	US-10-282-122A-50913	Sequence 50913, A
31	390	38.1	223	5	US-10-732-923-21104	Sequence 21104, A
32	371	36.2	194	5	US-10-732-923-21229	Sequence 21229, A
33	361.5	35.3	180	4	US-10-369-493-17716	Sequence 17716, A
34	361.5	35.3	180	5	US-10-732-923-21084	Sequence 21084, A
35	360	35.2	199	5	US-10-732-923-21100	Sequence 21100, A
36	358.5	35.0	200	5	US-10-732-923-21268	Sequence 21268, A
37	353.5	34.5	193	4	US-10-369-493-20936	Sequence 20936, A
38	353.5	34.5	200	5	US-10-732-923-21284	Sequence 21284, A
39	352	34.4	197	4	US-10-732-923-21042	Sequence 21042, A
40	351.5	34.3	192	4	US-10-369-493-18827	Sequence 18827, A
41	351.5	34.3	199	5	US-10-732-923-21049	Sequence 21049, A
42	351.5	34.3	203	5	US-10-732-923-20977	Sequence 20977, A
43	351	34.3	199	5	US-10-732-923-21152	Sequence 21152, A
44	351	34.3	199	5	US-10-732-923-21153	Sequence 21153, A
45	351	34.3	199	5	US-10-732-923-21206	Sequence 21206, A

ALIGNMENTS

RESULT 1
US-10-282-122A-62742

Sequence 62742, Application US/10282122A
Publication No. US20040029129A1

GENERAL INFORMATION:

APPLICANT: Wang, Liangu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zykied, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Foreyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

FILE REFERENCE: ELITRA.034A

CURRENT FILING DATE: 2003-02-20

PRIOR APPLICATION NUMBER: 60/191,078

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23

PRIOR APPLICATION NUMBER: 60/207,727

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/230,335

PRIOR FILING DATE: 2000-09-06

PRIOR APPLICATION NUMBER: 60/230,347

PRIOR FILING DATE: 2000-09-09

PRIOR APPLICATION NUMBER: 60/242,578

PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931

PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/267,636

PRIOR FILING DATE: 2001-02-09

PRIOR APPLICATION NUMBER: 60/269,308

PRIOR FILING DATE: 2001-02-16

Remaining Prior Application data removed - See File Wrapper or PALM.

NUMBER OF SEQ ID NOS: 78614

SOFTWARE: Patent version 3.1

SEQ ID NO 62742

LENGTH: 195

TYPE: PRT

ORGANISM: Mycobacterium bovis

US-10-282-122A-62742

Query Match 100.0%; Score 1024; DB 4; Length 195;
Beet Local Similarity 100.0%; Pred. No. 3.2e-103;
Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MPLLITGDOPFAVQYLTALIGDLSKVDKOPGDYFTTITSDHFGKRWVFPMPKDFTV 60
Db 1 MPLLITGDOPFAVQYLTALIGDLSKVDKOPGDYFTTITSDHFGKRWVFPMPKDFTV 60

Qy 61 CPTETIAAFSKLNDEFEDRDAQILGVSIDSEFAHFQWRAQHNDLKTLPFPMLSDIKRELISQ 120
Db 61 CPTETIAAFSKLNDEFEDRDAQILGVSIDSEFAHFQWRAQHNDLKTLPFPMLSDIKRELISQ 120

Qy 121 AAGVLANADGVADRVTFTVDPNNEIQFVSATAGSVGRNVDEVLRVLDALQSDDELCAQWNRK 180
Db 121 AAGVLANADGVADRVTFTVDPNNEIQFVSATAGSVGRNVDEVLRVLDALQSDDELCAQWNRK 180

Qy 181 GDDPTLDAGELLKASA 195
Db 181 GDDPTLDAGELLKASA 195

RESULT 2
US-10-282-122A-64708
; Sequence 64708, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Foreyeth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; PRIOR FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 64708
; LENGTH: 195
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-282-122A-64708

Query Match 100.0%; Score 1024; DB 4; Length 195;
Beet Local Similarity 100.0%; Pred. No. 3.2e-103;

Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MPLLITGDOPFAVQYLTALIGDLSKVDKOPGDYFTTITSDHFGKRWVFPMPKDFTV 60
Db 1 MPLLITGDOPFAVQYLTALIGDLSKVDKOPGDYFTTITSDHFGKRWVFPMPKDFTV 60

Qy 61 CPTETIAAFSKLNDEFEDRDAQILGVSIDSEFAHFQWRAQHNDLKTLPFPMLSDIKRELISQ 120
Db 61 CPTETIAAFSKLNDEFEDRDAQILGVSIDSEFAHFQWRAQHNDLKTLPFPMLSDIKRELISQ 120

Qy 121 AAGVLANADGVADRVTFTVDPNNEIQFVSATAGSVGRNVDEVLRVLDALQSDDELCAQWNRK 180
Db 121 AAGVLANADGVADRVTFTVDPNNEIQFVSATAGSVGRNVDEVLRVLDALQSDDELCAQWNRK 180

Qy 181 GDDPTLDAGELLKASA 195
Db 181 GDDPTLDAGELLKASA 195

RESULT 3
US-10-617-038-26
; Sequence 26, Application US/10617038
; Publication No. US20040057963A1
; GENERAL INFORMATION:
; APPLICANT: Andersen, Peter
; APPLICANT: Stryhn, Anette
; APPLICANT: Rosenkrands, Ida
; TITLE OF INVENTION: Therapeutic TB Vaccine
; FILE REFERENCE: S815AUSA
; CURRENT APPLICATION NUMBER: US/10/617,038
; PRIOR FILING DATE: 2003-07-11
; PRIOR APPLICATION NUMBER: DK PA 2002 01098
; PRIOR FILING DATE: 2002-07-13
; PRIOR APPLICATION NUMBER: US 60/401,725
; PRIOR FILING DATE: 2002-08-07
; NUMBER OF SEQ ID NOS: 187
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 26
; LENGTH: 195
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-617-038-26

Query Match 100.0%; Score 1024; DB 4; Length 195;
Beet Local Similarity 100.0%; Pred. No. 3.2e-103;
Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MPLLITGDOPFAVQYLTALIGDLSKVDKOPGDYFTTITSDHFGKRWVFPMPKDFTV 60
Db 1 MPLLITGDOPFAVQYLTALIGDLSKVDKOPGDYFTTITSDHFGKRWVFPMPKDFTV 60

Qy 61 CPTETIAAFSKLNDEFEDRDAQILGVSIDSEFAHFQWRAQHNDLKTLPFPMLSDIKRELISQ 120
Db 61 CPTETIAAFSKLNDEFEDRDAQILGVSIDSEFAHFQWRAQHNDLKTLPFPMLSDIKRELISQ 120

Qy 121 AAGVLANADGVADRVTFTVDPNNEIQFVSATAGSVGRNVDEVLRVLDALQSDDELCAQWNRK 180
Db 121 AAGVLANADGVADRVTFTVDPNNEIQFVSATAGSVGRNVDEVLRVLDALQSDDELCAQWNRK 180

Qy 181 GDDPTLDAGELLKASA 195
Db 181 GDDPTLDAGELLKASA 195

RESULT 4
US-10-332-512A-21
; Sequence 21, Application US/10332512A
; Publication No. US20040180056A1
; GENERAL INFORMATION:
; APPLICANT: ORME, Ian M.
; APPLICANT: BELISLE, John T.
; TITLE OF INVENTION: MID-LIFE VACCINE AND METHODS FOR BOOSTING ANTI-MYCOBACTERIAL IMMUN
; FILE REFERENCE: 38861-186292

```

; CURRENT APPLICATION NUMBER: US/10/332,512A
; CURRENT FILING DATE: 2003-01-10
; PRIOR APPLICATION NUMBER: PCT/US01/21717
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/217,646
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 21
; LENGTH: 195
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-332-512A-21
```

```

Query Match          100.0%; Score 1024; DB 4; Length 195;
Best Local Similarity 100.0%; Pred. No. 3.2e-103;
Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

Qy 1 MPLLTIQDPFAYQLTALIGDLSKVDAKQPGDYFTTITSDEHPGKMRVVFPMKDPFTFV 60
Db 1 MPLLTIQDPFAYQLTALIGDLSKVDAKQPGDYFTTITSDEHPGKMRVVFPMKDPFTFV 60
Qy 61 CPTETAAFSKLNDEFEDRDAQILGVSIIDSEFAHFQWRAQHNDLKTLPFPMLSDIKRELSQ 120
Db 61 CPTETAAFSKLNDEFEDRDAQILGVSIIDSEFAHFQWRAQHNDLKTLPFPMLSDIKRELSQ 120
Qy 121 AAGVLANADGVADRVTFTYIDPNNIEIOFVSATAGSVGRNVDEVLRVLDALQSDDELCAQNRK 180
Db 121 AAGVLANADGVADRVTFTYIDPNNIEIOFVSATAGSVGRNVDEVLRVLDALQSDDELCAQNRK 180
Qy 181 GDPTLDAGELLKASA 195
Db 181 GDPTLDAGELLKASA 195
```

```

RESULT 5
US-10-732-923-21281
; Sequence 21281, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 21281
; LENGTH: 195
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis H37Rv
US-10-732-923-21281
```

```

Query Match          100.0%; Score 1024; DB 5; Length 195;
Best Local Similarity 100.0%; Pred. No. 3.2e-103;
Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

Qy 1 MPLLTIQDPFAYQLTALIGDLSKVDAKQPGDYFTTITSDEHPGKMRVVFPMKDPFTFV 60
Db 1 MPLLTIQDPFAYQLTALIGDLSKVDAKQPGDYFTTITSDEHPGKMRVVFPMKDPFTFV 60
Qy 61 CPTETAAFSKLNDEFEDRDAQILGVSIIDSEFAHFQWRAQHNDLKTLPFPMLSDIKRELSQ 120
Db 61 CPTETAAFSKLNDEFEDRDAQILGVSIIDSEFAHFQWRAQHNDLKTLPFPMLSDIKRELSQ 120
Qy 121 AAGVLANADGVADRVTFTYIDPNNIEIOFVSATAGSVGRNVDEVLRVLDALQSDDELCAQNRK 180
Db 121 AAGVLANADGVADRVTFTYIDPNNIEIOFVSATAGSVGRNVDEVLRVLDALQSDDELCAQNRK 180
Qy 181 GDPTLDAGELLKASA 195
Db 181 GDPTLDAGELLKASA 195
```

```

RESULT 6
US-10-732-923-21020
; Sequence 21020, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 21020
; LENGTH: 195
; TYPE: PRT
; ORGANISM: Mycobacterium marinum
US-10-732-923-21020
```

```

Query Match          92.1%; Score 943; DB 5; Length 195;
Best Local Similarity 90.8%; Pred. No. 2.3e-94;
Matches 177; Conservative 7; Mismatches 11; Indels 0; Gaps 0;
```

```

Qy 1 MPLLTIQDPFAYQLTALIGDLSKVDAKQPGDYFTTITSDEHPGKMRVVFPMKDPFTFV 60
Db 1 MPLLTIQDPFAYQLTALIGDLSKVDAKQPGDYFTTITSDEHPGKMRVVFPMKDPFTFV 60
Qy 61 CPTETAAFSKLNDEFEDRDAQILGVSIIDSEFAHFQWRAQHNDLKTLPFPMLSDIKRELSQ 120
Db 61 CPTETAAFSKLNDEFEDRDAQILGVSIIDSEFAHFQWRAQHNDLKTLPFPMLSDIKRELSQ 120
Qy 121 AAGVLANADGVADRVTFTYIDPNNIEIOFVSATAGSVGRNVDEVLRVLDALQSDDELCAQNRK 180
Db 121 AAGVLANADGVADRVTFTYIDPNNIEIOFVSATAGSVGRNVDEVLRVLDALQSDDELCAQNRK 180
Qy 181 GDPTLDAGELLKASA 195
Db 181 GDPTLDAGELLKASA 195
```

```

RESULT 7
US-10-282-122A-61886
; Sequence 61886, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: EILTRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
```



```

; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See file Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 63991
; LENGTH: 195
; TYPE: PRT
; ORGANISM: Mycobacterium leprae
US-10-282-122A-63991
```

```

Query Match      88.3%; Score 904; DB 4; Length 195;
Best Local Similarity 86.7%; Pred. No. 4.1e-90;
Matches 169; Conservative 12; Mismatches 14; Indels 0; Gaps 0;
```

```

Qy      1 MFLITIGQFPAYQUTALIGDLSKVDAKOPGDYFTTITSDHFGKRWVFFWPKDFTFV 60
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      1 MSLISIGQFPAYQUTALIGDLSKVDAKOPGDYFTTVSSSHPKRWVFFWPKDFTFI 60
Qy      61 CPTBIAAFSKLNDEFEDRDAQILGVSIDSEFAHFQWRAQHNDLKTLPFPMLSDIKRELSQ 120
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      61 CPTBIAAFGKLNDEFEGGAQILGVSIDSEFVHFQWRAQHEDLKTLPFPMLSDIKRELSA 120
Qy      121 AAGVLNADGVADRVTFIVDPNNEIQFVSATAGSVGRNVDEVLRVLDALQSDCLCACMWRK 180
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      121 AAGVLNADGVADRVTFIVDPNNDIQFVSATAGSVGRNVEVLRVLDALQSDCLCACMWRK 180
Qy      181 GDPTLDAGELLKASA 195
      |||:|||||:
Db      181 GDPTLNATELLKTS 195
      |||:|||||:
```

```

RESULT 11
US-10-732-923-21017
; Sequence 21017, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 21017
; LENGTH: 195
; TYPE: PRT
; ORGANISM: Mycobacterium leprae
US-10-732-923-21017
```

```

Query Match      88.3%; Score 904; DB 5; Length 195;
Best Local Similarity 86.7%; Pred. No. 4.1e-90;
Matches 169; Conservative 12; Mismatches 14; Indels 0; Gaps 0;
```

```

Qy      1 MFLITIGQFPAYQUTALIGDLSKVDAKOPGDYFTTITSDHFGKRWVFFWPKDFTFV 60
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      1 MSLISIGQFPAYQUTALIGDLSKVDAKOPGDYFTTVSSSHPKRWVFFWPKDFTFI 60
Qy      61 CPTBIAAFSKLNDEFEDRDAQILGVSIDSEFAHFQWRAQHNDLKTLPFPMLSDIKRELSQ 120
```

```

Db      61 CPTBIAAFGKLNDEFEGGAQILGVSIDSEFVHFQWRAQHEDLKTLPFPMLSDIKRELSA 120
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy      121 AAGVLNADGVADRVTFIVDPNNEIQFVSATAGSVGRNVDEVLRVLDALQSDCLCACMWRK 180
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      121 AAGVLNADGVADRVTFIVDPNNDIQFVSATAGSVGRNVEVLRVLDALQSDCLCACMWRK 180
Qy      181 GDPTLDAGELLKASA 195
      |||:|||||:
Db      181 GDPTLNATELLKTS 195
      |||:|||||:
```

```

RESULT 12
US-10-732-923-21019
; Sequence 21019, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 21019
; LENGTH: 195
; TYPE: PRT
; ORGANISM: Mycobacterium smegmatis
US-10-732-923-21019
```

```

Query Match      85.7%; Score 878; DB 5; Length 195;
Best Local Similarity 82.2%; Pred. No. 2.8e-87;
Matches 157; Conservative 20; Mismatches 14; Indels 0; Gaps 0;
```

```

Qy      1 MFLITIGQFPAYQUTALIGDLSKVDAKOPGDYFTTITSDHFGKRWVFFWPKDFTFV 60
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      1 MALITIGDQPEYDITAVVGDSLKVDAKOPDDYFTVTSKDYEGKRWIIFFWPKDFTFV 60
Qy      61 CPTBIAAFSKLNDEFEDRDAQILGVSIDSEFAHFQWRAQHNDLKTLPFPMLSDIKRELSQ 120
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      61 CPTBIAAFGKLNDEFEDRDAKVLGVSVDNEVHFQWRAQHEDLKTLPFPMVSDIKRELSA 120
Qy      121 AAGVLNADGVADRVTFIVDPNNEIQFVSATAGSVGRNVDEVLRVLDALQSDCLCACMWRK 180
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      121 AAGVLNADGVADRVTFIVDPNNEIQFVSATAGSVGRNVEVLRVLDALQSDCLCACMWRK 180
Qy      181 GDPTLDAGELL 191
      |||:|||||:
Db      181 GDPTINAGELL 191
      |||:|||||:
```

```

RESULT 13
US-10-282-122A-53752
; Sequence 53752, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zykkind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Cary, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: EITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
```


GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OW protein - protein search, using SW model

Run on: March 23, 2006, 06:17:23 ; Search time 4.67951 Seconds
(without alignments)
1192.746 Million cell updates/sec

Title: US-10-617-038-26

Perfect score: 1024
Sequence: 1 MPLLTIGDQFPAYQLTALIG.....CNMRKGPPTLDAGEILKASA 195

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 169630 seqs, 28622889 residues

Total number of hits satisfying chosen parameters: 169630

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_AA_New:*
1: /cgn2_6/ptodata/1/pubpaa/US06_NEM_PUB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/US06_NEM_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US07_NEM_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/BCT_NEM_PUB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US08_NEM_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US10_NEM_PUB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US11_NEM_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US60_NEM_PUB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	350.5	34.2	446	US-11-108-172-1121	Sequence 1121, Ap
2	318	31.1	206	US-10-821-234-1038	Sequence 1038, Ap
3	318	31.1	257	US-10-642-272A-6	Sequence 6, Appl
4	317	31.0	256	US-10-642-272A-4	Sequence 4, Appl
5	317	31.0	257	US-10-642-272A-5	Sequence 5, Appl
6	312	30.5	198	US-10-642-272A-28	Sequence 28, Appl
7	298	29.1	193	US-10-527-771-6	Sequence 6, Appl
8	274	26.8	187	US-11-058-926-22	Sequence 22, Appl
9	271.5	26.5	271	US-10-821-234-1419	Sequence 1419, Ap
10	263.5	25.7	188	US-11-058-926-24	Sequence 24, Appl
11	261.5	25.5	187	US-11-077-619-4	Sequence 4, Appl
12	259.5	25.3	187	US-11-077-619-86	Sequence 86, Appl
13	210.5	20.6	216	US-11-096-568A-33952	Sequence 33952, A
14	210.5	20.6	229	US-11-096-568A-657	Sequence 657, App
15	210.5	20.6	245	US-11-096-568A-656	Sequence 656, App
16	210.5	20.6	252	US-11-096-568A-33951	Sequence 33951, A
17	142.5	13.9	216	US-11-096-568A-1037	Sequence 1037, Ap
18	142.5	13.9	226	US-11-096-568A-1038	Sequence 1038, Ap
19	140	13.7	180	US-11-096-568A-658	Sequence 658, App
20	133	13.0	218	US-11-096-568A-21440	Sequence 21440, A
21	128	12.5	164	US-11-058-926-14	Sequence 33953, A
22	103	10.1	32	US-11-058-926-14	Sequence 19, Appl
23	102	10.0	32	US-11-058-926-14	Sequence 14, Appl
24	96.5	9.4	167	US-10-467-657-2342	Sequence 2942, Ap
25	95	9.3	32	US-11-058-926-20	Sequence 20, Appl

26	92.5	9.0	32	7	US-11-058-926-12	Sequence 12, Appl
27	92.5	9.0	32	7	US-11-058-926-13	Sequence 13, Appl
28	92	9.0	164	6	US-10-793-626-252	Sequence 252, App
29	91	8.9	32	7	US-11-058-926-18	Sequence 18, Appl
30	85.5	8.3	32	7	US-11-058-926-17	Sequence 17, Appl
31	83	8.1	164	7	US-11-074-176-70	Sequence 70, Appl
32	82	8.0	376	7	US-11-055-822-422	Sequence 422, App
33	82	8.0	518	7	US-11-055-822-420	Sequence 420, App
34	81	7.9	32	7	US-11-058-926-15	Sequence 15, Appl
35	80	7.8	162	6	US-10-467-657-1176	Sequence 1176, Ap
36	77.5	7.6	168	7	US-11-087-099-10829	Sequence 10829, A
37	77.5	7.6	3361	6	US-10-453-372-1082	Sequence 1082, Ap
38	77	7.5	32	7	US-11-058-926-16	Sequence 16, Appl
39	77	7.5	200	7	US-11-087-099-4145	Sequence 4145, Ap
40	77	7.5	307	7	US-11-096-568A-20009	Sequence 20009, A
41	77	7.5	348	7	US-11-096-568A-20008	Sequence 17, Appl
42	76	7.4	15	7	US-11-058-926-9	Sequence 9, Appl
43	76	7.4	1307	7	US-11-072-512-2438	Sequence 2438, Ap
44	74	7.2	230	7	US-11-087-099-5639	Sequence 5639, Ap
45	74	7.2	231	7	US-11-087-099-5149	Sequence 9149, Ap

ALIGNMENTS

RESULT 1
US-11-108-172-1121
Sequence 1121, Application US/11108172
Publication No. US20050260177A1
GENERAL INFORMATION:
APPLICANT: Xu, Jianshun
APPLICANT: Lodes, Michael J.
APPLICANT: Secretist, Heather
APPLICANT: Benson, Darin R.
APPLICANT: Meagher, Madeleine Joy
APPLICANT: Stoik, John A.
APPLICANT: Wang, Tongtong
APPLICANT: Jiansun, Yugu
APPLICANT: Smith, Carole L.
APPLICANT: King, Gordon E.
APPLICANT: Wang, Aijun
APPLICANT: Clapper, Jonathan D.
APPLICANT: Skeiky, Yasir A. W.
APPLICANT: Fanger, Gary R.
APPLICANT: Vedvick Thomas S.
APPLICANT: Carter, Darick
TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
TITLE OF INVENTION: OF COLON CANCER AND METHODS FOR THEIR USE
FILE REFERENCE: 210121.471C15
CURRENT APPLICATION NUMBER: US/11/108, 172
CURRENT FILING DATE: 2005-04-15
PRIOR APPLICATION NUMBER: US 10/025,380
PRIOR FILING DATE: 2001-12-19
PRIOR APPLICATION NUMBER: US 09/922,217
PRIOR FILING DATE: 2001-08-03
PRIOR APPLICATION NUMBER: US 09/833,263
PRIOR FILING DATE: 2001-04-10
PRIOR APPLICATION NUMBER: US 09/649,811
PRIOR FILING DATE: 2000-08-28
PRIOR APPLICATION NUMBER: US 09/609,448
PRIOR FILING DATE: 2000-06-29
PRIOR APPLICATION NUMBER: US 09/575,251
PRIOR FILING DATE: 2000-05-19
PRIOR APPLICATION NUMBER: US 09/519,444
PRIOR FILING DATE: 2000-03-06
PRIOR APPLICATION NUMBER: US 09/504,629
PRIOR FILING DATE: 2000-02-15
PRIOR APPLICATION NUMBER: US 09/480,321
PRIOR FILING DATE: 2000-01-10
PRIOR APPLICATION NUMBER: US 09/476,296
PRIOR FILING DATE: 1999-12-30
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 1130

```
; SOFTWARE: FactSeq for Windows Version 4.0
; SEQ ID NO 1121
; LENGTH: 446
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-108-172-1121
```

```
Query Match          34.2%; Score 350.5; DB 7; Length 446;
Best Local Similarity 39.4%; Pred. No. 5,2e-28;
Matches 74; Conservative 35; Mismatches 60; Indels 19; Gaps 4;
```

```
Qy 6 IGDPAYQVLTALIGDLSKVDAKQPGDYFTTTSDEHFGKRVVFPWKDFTVCPTPEI 65
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 8 INSPAPSEEVALLM-----PNSGFKKISLSYKGVVLFYPLDFTVCPTPEV 56
Qy 66 AAFSKLNDDEFDRDAQILGVSIDSEFAHFWRAQ---HNDKLTLPFPMLSDIKRELQQA 122
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 57 IAFSGSVSRFBNELNEVLACGIDSEYALQWLTQDRKKGGGLCTMAIPMLADKTSIASY 116
Qy 123 GVLNADGVADRVTFIVDPNNEIQFVSATAGSVGNVDEVLRVLDALQ---SDELCA 177
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 117 GVLSESGVAVRGLFIIDPHGKLRQITVNDMPVGRSVEVLRLLEAFQFVEKHGEVCPAN 176
Qy 178 WRKGDPTL 185
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 177 WKKGAPTM 184
```

```
RESULT 2
US-10-821-234-1038
; Sequence 1038, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; PRIOR FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pc_seq_genes Version 1.0
; SEQ ID NO 1038
; LENGTH: 206
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-821-234-1038
```

```
Query Match          31.1%; Score 318; DB 6; Length 206;
Best Local Similarity 36.7%; Pred. No. 3,9e-25;
Matches 69; Conservative 33; Mismatches 66; Indels 20; Gaps 4;
```

```
Qy 6 IGDPAYQVLTALIGDLSKVDAKQPGDYFTTTSDEHFGKRVVFPWKDFTVCPTPEI 65
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 16 IGRDAPFKATVAVDGD-----AFKEVKLSYKGVVLFYPLDFTVCPTPEI 63
Qy 66 AAFSKLNDDEFDRDAQILGVSIDSEFAHFWRAQ---HNDKLTLPFPMLSDIKRELQQA 122
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 64 IAFSNRAEDFRKGLGCEVIGSVSDQFTHLAMINTPRKEGGGAPNIPPLADVTRELSBDY 123
Qy 123 GVLNADGVADRVTFIVDPNNEIQFVSATAGSVGNVDEVLRVLDALQ---SDELCA 177
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 124 GVLKTDGIAVRGLFIIDPHGKLRQITVNDLPVGRSVEVLRLLEAFQFVEKHGEVCPAN 183
Qy 178 WRKGDPTL 185
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 184 WKPGSDTI 191
```

```
RESULT 3
```

```
US-10-642-272A-6
; Sequence 6, Application US/10642272A
; Publication No. US20050277606A1
; GENERAL INFORMATION:
; APPLICANT: Hattori, Fumiyuki
; APPLICANT: Sugimura, Kei-jiro
; APPLICANT: Furuya, Mayumi
; TITLE OF INVENTION: Therapeutic Methods and Agents for Diseases Associated with
; FILE REFERENCE: 58777.000012
; CURRENT APPLICATION NUMBER: US/10/642,272A
; PRIOR APPLICATION NUMBER: PCT/JP02/01358
; PRIOR FILING DATE: 2001-02-18
; PRIOR APPLICATION NUMBER: JP 41003/2001
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 6
; LENGTH: 257
; TYPE: PRT
; ORGANISM: Mus sp.
US-10-642-272A-6
```

```
Query Match          31.1%; Score 318; DB 6; Length 257;
Best Local Similarity 37.2%; Pred. No. 5,3e-25;
Matches 68; Conservative 38; Mismatches 57; Indels 20; Gaps 4;
```

```
Qy 11 PAYQVLTALIGDLSKVDAKQPGDYFTTTSDEHFGKRVVFPWKDFTVCPTPEIAFSK 70
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 71 PYFKGIVAVNGE-----FKELSDDRKGLVLFYPLDFTVCPTPEIYVAFSD 118
Qy 71 LNDEFDRDAQILGVSIDSEFAHFWRAQ---HNDKLTLPFPMLSDIKRELQQAQV-LN 126
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 119 KANEFHVNCEVAVASVDSHFSHLAMINTPRKKGGLGHNMITLSDITKQISHDYGLLE 178
Qy 127 ADGVADRVTFIVDPNNEIQFVSATAGSVGNVDEVLRVLDALQ---SDELCA 182
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 179 SAGIALRGFLIIDPHGKLRQITVNDLPVGRSVEVLRLLEAFQFVEKHGEVCPANWPES 238
Qy 183 PTL 185
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 239 PTI 241
```

```
RESULT 4
US-10-642-272A-4
; Sequence 4, Application US/10642272A
; Publication No. US20050277606A1
; GENERAL INFORMATION:
; APPLICANT: Hattori, Fumiyuki
; APPLICANT: Sugimura, Kei-jiro
; APPLICANT: Furuya, Mayumi
; TITLE OF INVENTION: Therapeutic Methods and Agents for Diseases Associated with
; FILE REFERENCE: 58777.000012
; CURRENT APPLICATION NUMBER: US/10/642,272A
; PRIOR APPLICATION NUMBER: PCT/JP02/01358
; PRIOR FILING DATE: 2001-02-18
; PRIOR APPLICATION NUMBER: JP 41003/2001
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 4
; LENGTH: 256
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-642-272A-4
```

```
Query Match          31.0%; Score 317; DB 6; Length 256;
Best Local Similarity 37.2%; Pred. No. 6,6e-25;
Matches 68; Conservative 37; Mismatches 58; Indels 20; Gaps 4;
```


Qy 123 GVLAAD-GVADRVTFIVDPNNEIQFVSATAGSVGRNDEVLRVLDALQ----SDELCAKN 177
Db 111 GVLEDEBIAVRGGLFTIDPKGILRQITVNDLPVGRSVDETLRLVQAFQYVDKRGVCPAG 170
Qy 178 WRKGDPTLDAG 188
Db 171 WTPGKATIKPG 181

RESULT 8

US-11-058-926-22
; Sequence 22, Application US/11058926
; Publication No. US20060030022A1
; GENERAL INFORMATION:
; APPLICANT: BECKWITH, JONATHAN
; APPLICANT: ASLUND, FREDRIK
; APPLICANT: BESSETTE, PAUL H.
; APPLICANT: GEORGIU, GEORGE
; APPLICANT: RITZ, DANIEL
; APPLICANT: LIM, JACKIE EUN-AH
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR PRODUCTION OF DISULFIDE
; FILE REFERENCE: HMV-052.01
; CURRENT APPLICATION NUMBER: US/11/058,926
; CURRENT FILING DATE: 2005-02-16
; PRIOR APPLICATION NUMBER: US/09/679,705
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/157,770
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/163,939
; PRIOR FILING DATE: 1999-11-08
; PRIOR APPLICATION NUMBER: 60/166,044
; PRIOR FILING DATE: 1999-11-17
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 22
; LENGTH: 187
; TYPE: PRT
; ORGANISM: Escherichia coli
US-11-058-926-22

Query Match 26.8%; Score 274; DB 7; Length 187;
Best Local Similarity 34.4%; Pred. No. 1.1e-20;
Matches 54; Conservative 29; Mismatches 68; Indels 6; Gaps 2;
Qy 35 FTTTSDHFGKRV-VFWPMDFTVCPTETIAAFSKLNDEFEDRDAQILGVSIDSEFAHF 94
Db 21 FIETKDETRGRMSVFFFPADFTVCPTETELGDVAHDELOKLGVDVYAVSTDTHTFH 80
Qy 95 QWRAGHNDLKTLPFPMLSIDIRELSQAAGVLNAD-GVADRVTFIVDPNNEIQFVSATAGS 153
Db 81 AMHSSSETIAKIKYAMIGDPTGALTFRNFDNMRDEGLADRAFTFVDDPGIIGALEVTAEG 140
Qy 154 VGRVDEVLRVLDALQ-----SDELCAKNMRKGDPTL 185
Db 141 IGRDASDLRKIKAAQYVASHPEVCAPKMRGEATL 177

RESULT 9

US-10-821-234-1419
; Sequence 1419, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Blythe
; APPLICANT: Andaman, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07

; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: PRT-Seq, Genes Version 1.0
; SEQ ID NO 1419
; LENGTH: 271
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-821-234-1419

Query Match 26.5%; Score 271.5; DB 6; Length 271;
Best Local Similarity 33.9%; Pred. No. 3.2e-20;
Matches 64; Conservative 32; Mismatches 72; Indels 21; Gaps 5;

Qy 23 LSKVDARQPDY-----FTTTSDEHFGKRV-VFWPMDFTVCPTETIAAFSKLN 72
Db 76 LSKAKISKPAWEGTAVIDGEPKELTLYRGKYLVFFFPDFTVCPTETIAAFGDL 135
Qy 73 DEFEDRDAQILGVSIDSEFAHFQW--RAOHNDLKTLPFPMLSIDIRELSQAAGV-LNAD 128
Db 136 EEFSSINTEVYVAGSVDSQFTHLAMINTPRQGLGPRIRILSDTLTIQISKDYGVLEDS 195
Qy 129 GVADRVTFIVDPNNEIQFVSATAGSVGRNDEVLRVLDALQ----SDELCAKNMRKGDPT 184
Db 196 GHTLRGLFIIDDKILRQITLNDLPVGRSVDETLRLVQAFQYTDKRGVCPAGKPGSET 255
Qy 185 L--DAGEL 190
Db 256 IIPDPAGKL 264

RESULT 10

US-11-058-926-24
; Sequence 24, Application US/11058926
; Publication No. US20060030022A1
; GENERAL INFORMATION:
; APPLICANT: BECKWITH, JONATHAN
; APPLICANT: ASLUND, FREDRIK
; APPLICANT: BESSETTE, PAUL H.
; APPLICANT: GEORGIU, GEORGE
; APPLICANT: RITZ, DANIEL
; APPLICANT: LIM, JACKIE EUN-AH
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR PRODUCTION OF DISULFIDE
; FILE REFERENCE: HMV-052.01
; CURRENT APPLICATION NUMBER: US/11/058,926
; CURRENT FILING DATE: 2005-02-16
; PRIOR APPLICATION NUMBER: US/09/679,705
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/157,770
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/163,939
; PRIOR FILING DATE: 1999-11-08
; PRIOR APPLICATION NUMBER: 60/166,044
; PRIOR FILING DATE: 1999-11-17
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 24
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Escherichia coli
US-11-058-926-24

Query Match 25.7%; Score 263.5; DB 7; Length 188;
Best Local Similarity 34.2%; Pred. No. 1.3e-19;
Matches 54; Conservative 29; Mismatches 68; Indels 7; Gaps 3;
Qy 35 FTTTSDHFGKRV-VFWPMDFTVCPTETIAAFSKLNDEFEDRDAQILGVSIDSEFAH 93
Db 21 FIETKDETRGRMSVFFFPADFTVCPTETELGDVAHDELOKLGVDVYAVSTDTHTFH 80
Qy 94 QWRAGHNDLKTLPFPMLSIDIRELSQAAGVLNAD-GVADRVTFIVDPNNEIQFVSATAG 152
Db 81 KAHSSSETIAKIKYAMIGDPTGALTFRNFDNMRDEGLADRAFTFVDDPGIIGALEVTA 140

Qy 153 SVGRNVDVLRVLDALQ-----SDELCACNMRKGDPTL 185
Db 141 GIGRDSADLLRKIKAAQYVASHPGSEVCPAKWEGEATL 178

RESULT 11

US-11-077-619-4
Sequence 4, Application US/11077619
Publication No. US20060040279A1
GENERAL INFORMATION:
APPLICANT: Feesche, Joerg
APPLICANT: Maurer, Karl-Heinz
APPLICANT: Breves, Roland
APPLICANT: Schneider, Thomas
APPLICANT: Hecker, Michael
APPLICANT: Juergen, Britta
APPLICANT: Voigt, Birgit
TITLE OF INVENTION: DNA CHIPS USED FOR BIOPROCESS CONTROL
FILE REFERENCE: HENK-0122 / H5692
CURRENT APPLICATION NUMBER: US/11/077,619
PRIOR FILING DATE: 2005-03-11
PRIOR APPLICATION NUMBER: PCT/EP2003/009979
PRIOR FILING DATE: 2003-09-09
PRIOR APPLICATION NUMBER: DE 10242433.0
PRIOR FILING DATE: 2002-09-11
NUMBER OF SEQ ID NOS: 130
SOFTWARE: PatentIn version 3.3
SEQ ID NO 4
LENGTH: 187
TYPE: PRT
ORGANISM: Bacillus subtilis
US-11-077-619-4

Query Match 25.5%; Score 261.5; DB 7; Length 187;
Best Local Similarity 31.6%; Pred. No. 2.1e-19;
Matches 56; Conservative 39; Mismatches 73; Indels 9; Gaps 4;

Qy 17 ALIGDLSKVDAK--QPGDYFTTTSDEHPGKRVVFPMPKDFVCPTEIAAFSKLND 74
Db 2 SLIGKEVPPFAKAKNGE-FIDVTNEDLKQMSFCFYPADFSVCTELEDLOEQYAA 60
Qy 75 FEDRDAQILGVSIDSEFAHFQWRAQHNDLKTLPFPMLSDIKRELSQAAGVLAN-ADGVA 133
Db 61 LKEIGVEYVSSTDTHTVHKGMHDSSEKISKITYAMIGDPSQGISRNPDVLDDEETGLADR 120
Qy 134 VTPIVDPNNEIQFVSATGSGVGRNVDEVLRVLDALQ-----SDELCACNMRKGDPTL 185
Db 121 GTFIIDPDGVIOQAVEINAGGIGRDASNLVNVKKAQYVRONPGSEVCPAKWEGEATL 177

RESULT 12

US-11-077-619-86
Sequence 86, Application US/11077619
Publication No. US20060040279A1
GENERAL INFORMATION:
APPLICANT: Feesche, Joerg
APPLICANT: Maurer, Karl-Heinz
APPLICANT: Breves, Roland
APPLICANT: Schneider, Thomas
APPLICANT: Hecker, Michael
APPLICANT: Juergen, Britta
APPLICANT: Voigt, Birgit
TITLE OF INVENTION: DNA CHIPS USED FOR BIOPROCESS CONTROL
FILE REFERENCE: HENK-0122 / H5692
CURRENT APPLICATION NUMBER: US/11/077,619
PRIOR FILING DATE: 2005-03-11
PRIOR APPLICATION NUMBER: PCT/EP2003/009979
PRIOR FILING DATE: 2003-09-09
PRIOR APPLICATION NUMBER: DE 10242433.0
PRIOR FILING DATE: 2002-09-11
NUMBER OF SEQ ID NOS: 130
SOFTWARE: PatentIn version 3.3
SEQ ID NO 86

LENGTH: 187
TYPE: PRT
ORGANISM: Bacillus licheniformis
US-11-077-619-86

Query Match 25.3%; Score 259.5; DB 7; Length 187;
Best Local Similarity 31.6%; Pred. No. 3.3e-19;
Matches 56; Conservative 37; Mismatches 75; Indels 9; Gaps 4;

Qy 17 ALIGDLSKVDAK--QPGDYFTTTSDEHPGKRVVFPMPKDFVCPTEIAAFSKLND 74
Db 2 SLIGKEVPPFAKAKNGE-FIDVTNEDLKQMSFCFYPADFSVCTELEDLOEQYAA 60
Qy 75 FEDRDAQILGVSIDSEFAHFQWRAQHNDLKTLPFPMLSDIKRELSQAAGVLAN-ADGVA 133
Db 61 LKEIGVEYVSSTDTHTVHKGMHDSSEKISKITYAMIGDPSQGISRNPDVLDDEETGLADR 120
Qy 134 VTPIVDPNNEIQFVSATGSGVGRNVDEVLRVLDALQ-----SDELCACNMRKGDPTL 185
Db 121 GTFIIDPDGVIOQAVEINAGGIGRDASILNVKKAQYVRONPGSEVCPAKWEGEATL 177

RESULT 13

US-11-096-568A-33952
Sequence 33952, Application US/11096568A
Publication No. US20060048240A1
GENERAL INFORMATION:
APPLICANT: Alexandrov, Nickolai et al.
TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
FILE REFERENCE: 2750-1592PUS2
CURRENT APPLICATION NUMBER: US/11/096,568A
CURRENT FILING DATE: 2005-04-01
NUMBER OF SEQ ID NOS: 34471
SEQ ID NO 33952
LENGTH: 216
TYPE: PRT
ORGANISM: Arabidopsis thaliana
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)..(216)
OTHER INFORMATION: Cereals Seq. ID no. 13604610
US-11-096-568A-33952

Query Match 20.6%; Score 210.5; DB 7; Length 216;
Best Local Similarity 28.2%; Pred. No. 4.1e-14;
Matches 55; Conservative 38; Mismatches 77; Indels 25; Gaps 6;

Qy 1 MPILTTIGDQFPAYQUTALALIGDLSKVDAKQPGDYFTTTSDEHPGKRVVFPMPKDFPV 60
Db 1 MPGLITLGDVTPNLEVE-----TTHDKFKLHDVFA-----NSMTVLFSHGDFTPV 45
Qy 61 CPTETIAAFSKINDEFEDRDAQILGVSIDSEFAHFQWRAQ--HNDLKTLPFPMLSDIKRE 117
Db 46 CTTETLGAAMAKYAHBFDRKGVLTGLSCDVQSHKDWIKDLEAFHSGSKVVPPIIADNKE 105
Qy 118 LSGAAGVLAN-ADGVADEVTPITVDPNNEIQFVSATGSGVGRNVDEVLRVLDAL-----QSD 171
Db 106 IIPQANMIDPLENGSPALHIVGDSKIKSLFYPTTGKRWDEVLRALDSLMAASKHN 165
Qy 172 ELCA-CNMRKGDPTL 185
Db 166 KIATPVMKKDPQPV 180

RESULT 14

US-11-096-568A-657
Sequence 657, Application US/11096568A
Publication No. US20060048240A1
GENERAL INFORMATION:
APPLICANT: Alexandrov, Nickolai et al.
TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
FILE REFERENCE: 2750-1592PUS2
CURRENT APPLICATION NUMBER: US/11/096,568A
CURRENT FILING DATE: 2005-04-01
NUMBER OF SEQ ID NOS: 34471
SEQ ID NO 657

```

: FILE REFERENCE: 2750-1592PUS2
: CURRENT APPLICATION NUMBER: US/11/096,568A
: CURRENT FILING DATE: 2005-04-01
: NUMBER OF SEQ ID NOS: 34471
: SEQ ID NO 657
: LENGTH: 229
: TYPE: PRT
: ORGANISM: Zea mays subsp. mays
: FEATURE:
: NAME/KEY: misc.feature
: LOCATION: (1) -- (229)
: OTHER INFORMATION: Ceres Seq. ID no. 12637707
: US-11-096-568A-657

```

Query Match	20.6%;	Score 210.5;	DB 7;	Length 229;
Best Local Similarity	27.5%;	Pred. No. 4.5e-14;		
Matches 57;	Conservative 34;	Mismatches 75;	Indels 41;	Gaps 5;

```

QY      MPELTIGDPRPALGALGSLKVDKAPQPDYFTTISDHNPK -RRVVFYRKPQET 59
Db      1 MPELTIGDYPVNLIELSTHG-----KIRINDYGDGYAII FSHPADFT 44

QY      60 VCEETIAAFSKLNDPEFDRDAQIGVSLIDSEFAHFQW-----RAGHNDLKTLP 108
Db      45 VCTTEMAAAGYAKKEKRGVKLLGIGSCDDVESHQWTKDVEAYGKQKQQAATTKVF 104

QY      109 PMLSDIKRELISQAAGVLND-----GVADRVTFIVDPNNELQFVSATAGSVGRNDEV 161
Db      105 PTLADPARDAIRQLNANVDDEKDDAAGRSWPSALHVVGPDKAVKLSFLYPATTRNMDEV 164

QY      162 LRVLDALQS-----DELCAQNMKRC 182
Db      165 LRAVDSILTPAAKHGKGVATPANNMKPE 191

```

```

RESULT 15
US-11-096-568A-656
: Sequence 656, Application US/11096568A
: Publication No. US20060048240A1
: GENERAL INFORMATION:
: APPLICANT: Alexandrov, Nikolai et al.
: TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
: TITLE OF INVENTION: Theory
: FILE REFERENCE: 2750-1592PUS2
: CURRENT APPLICATION NUMBER: US/11/096,568A
: CURRENT FILING DATE: 2005-04-01
: NUMBER OF SEQ ID NOS: 34471
: SEQ ID NO 656
: LENGTH: 245
: TYPE: PRT
: ORGANISM: Zea mays subsp. mays
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (1)..(245)
: OTHER INFORMATION: Cerec Seq. ID no. 12637706
: US-11-096-568A-656

```

Query Match	20.6%;	Score 210.5;	DB 7;	Length 245;
Best Local Similarity	27.5%;	Pred. No. 4.9e-14;		
Matches 57;	Conservative 34;	Mismatches 75;	Indels 41;	Gaps 5;

```

QY 2 MELLTGGDGPAYQVLTALIGSLSKVDKAKQEDYTTTITSDERHGK -WRVVFPPMDFTF 55
D 17 MPELTGTGDVVPNLELDSTHG-----KLRHDDYGDGVAII FSHPADFTF 60
QY 60 VCPTEIARFSKUNDEFERDQAIIIGVSDSEPAHQW-----RAQHDNKLTPF 108
D 61 VCTTEMAANAAGYAKERKRVKLLGISCDVESHQWTKDVEAYGKQKQDQAATTKYTF 120
QY 109 PMLSDIKRELSQAAGVLAND-----GVADRVTFPIYDNPNNIGFVSATAGSYGRANDVEY 161
D 121 PIIADBARARAIQIQLNNVDPDEKDAAGRSMPSPALHVVGPDKAVKLSFLYPATGTGRNNDEV 180

```

```

QY      162 LRVLDALQS-----DELCACMNRKGD 182
          ||  ::||  ||  ::||
DB      181 LRAVDSLITAKHGKQVATPANWKPE 207

```

Search completed: March 23, 2006, 06:29:37
Job time : 5.67951 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: March 23, 2006, 04:56:29 ; Search time 72.3158 Seconds
(without alignments)
1652.628 Million cell updates/sec

Title: US-10-617-038-27

Perfect score: 1337
Sequence: 1 MSGSGEPMTKTIIVGIDSGH.....GPSGHFVRHAECSVLVVR 272

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21:*

1:	geneeqp1980s:*
2:	geneeqp1980s:*
3:	geneeqp2000s:*
4:	geneeqp2001s:*
5:	geneeqp2002s:*
6:	geneeqp2003as:*
7:	geneeqp2003bs:*
8:	geneeqp2004s:*
9:	geneeqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1397	100.0	272	8	AD137306 M. tuberc
2	416	29.8	279	8	AD137300 M. tuberc
3	357.5	25.6	268	8	AD137321 M. tuberc
4	286	20.5	295	8	AD137298 M. tuberc
5	267	19.1	297	5	AY87878 M. tuberc
6	267	19.1	297	5	ABG30907 Mycobacte
7	266	19.0	301	4	AB79640 Corynebact
8	266	19.0	301	4	AAU71872 C. glutam
9	266	19.0	301	4	AAU71865 C. glutam
10	266	19.0	301	4	AAU71865 C. glutam
11	266	19.0	301	5	ABG80324 C. glutam
12	266	19.0	301	9	ABE15399 C. glutam
13	249	17.8	300	6	ABU25937 Protein e
14	151	10.8	316	4	AAU62080 Propionib
15	151	10.8	316	6	ABM58599 Propionib
16	128.5	9.2	173	8	ADU05947 Plant ful
17	127	9.1	179	6	ADA48734 Rice prot
18	127	9.1	179	7	ABM90364 Rice prot
19	124	8.9	165	7	ADA48322 Rice prot
20	124	8.9	165	7	ABM86830 Rice abio
21	124	8.9	165	9	ADY60949 Abiotic s
22	121.5	8.7	524	2	AAW34202 Streptomy
23	121.5	8.7	524	2	AAW55803 Streptomy
24	116.5	8.3	162	3	AAU42273 Arabidops

25	116.5	8.3	198	3	AAU42272 Arabidops
26	113	8.1	153	3	AAU42274 Arabidops
27	112.5	8.1	162	3	AAU42274 Arabidops
28	112.5	8.1	183	3	AAU42274 Arabidops
29	112.5	8.0	189	5	ABP38950 Staphyloc
30	112	8.0	189	5	ABP38950 Staphyloc
31	110	7.9	149	7	ADM25915 Hypertner
32	109.5	7.8	166	6	ABM73534 Staphyloc
33	109	7.8	153	3	AAU54318 Arabidops
34	109	7.8	158	7	ADM25536 Hypertner
35	109	7.8	292	5	ABP07225 Human ORF
36	109	7.8	292	5	ABP07225 Human ORF
37	109	7.8	941	6	ABU36688 Protein e
38	109	7.8	941	6	ABU34503 Protein e
39	108	7.7	147	4	AAU91259 C. glutam
40	108	7.7	245	8	ADY06981 Plant ful
41	107.5	7.7	6751	8	ADU91702 Polyketid
42	107.5	7.7	6751	8	ABE86594 Streptomy
43	107.5	7.7	6751	8	ABE86594 Streptomy
44	107	7.7	206	8	ADU77663 Plant ful
45	105.5	7.6	613	4	AAU04887 Micromono

ALIGNMENTS

RESULT 1					
AD137306	AD137306	standard; protein; 272 AA.			
XX	AD137306;				
AC	AD137306;				
XX	AD137306;				
DT	22-APR-2004	(first entry)			
XX	22-APR-2004				
DE	M. tuberculosis low oxygen induced antigen Rv2624c SEQ ID NO:27.				
XX	M. tuberculosis low oxygen induced antigen; vaccine; tuberculosis;				
KW	Mycobacterium tuberculosis; immunisation; antibacterial; gene therapy;				
KW	low oxygen induced antigen.				
XX					
OS	Mycobacterium tuberculosis.				
XX					
PN	WO2004006952-A2.				
XX					
PD	22-JAN-2004.				
XX					
PF	08-JUL-2003; 2003WO-DK000477.				
XX					
PR	13-JUL-2002; 2002DK-00001098.				
XX					
PA	(STAT-) STATENS SERUM INST.				
XX					
PI	Andersen P, Rosenkrands I, Stryhn A;				
XX	WPI, 2004-122778/12.				
DR	N-PSDB; AD137351.				
XX					
PT	Use of one or more polypeptides or their fragments, which are expressed				
PT	during the latent stage of the mycobacterial infection, and/or nucleic				
PT	acids encoding the polypeptides, for a therapeutic vaccine against				
PT	tuberculosis.				
XX					
PS	Claim 3; SEQ ID NO 27; 76pp; English.				
XX					
CC	The present invention describes polypeptides or their fragments, which				
CC	are expressed during the latent stage of a mycobacterial infection,				
CC	and/or nucleic acids encoding the polypeptides, which are useful for				
CC	creating a therapeutic vaccine against tuberculosis. Also described: (1)				
CC	a therapeutic vaccine against tuberculosis comprising one or more				
CC	polypeptides; (2) a method for treating an animal, including a human				
CC	being, with tuberculosis caused by virulent mycobacteria, e.g. by				
CC	Mycobacterium tuberculosis, M. africanum or M. bovis; (3) a method for				
CC	immunising an animal, including a human being, against tuberculosis				

CC caused by virulent mycobacteria; (4) a method of diagnosing tuberculosis
CC caused by virulent mycobacteria in an animal, including a human being;
CC (5) a method for diagnosing previous or ongoing infection with a virulent
CC mycobacterium; and (6) a method of diagnosing Mycobacterium tuberculosis
CC infection in a subject. The polypeptides have antibacterial activities,
CC and can be used in vaccines and in gene therapy. The polypeptides are
CC useful for the manufacture of a therapeutic vaccine for treating an
CC individual who is infected by a virulent mycobacterium, e.g. M.
CC tuberculosis, and who is not vaccinated with BCG against tuberculosis.
CC The present sequence represents a low oxygen induced antigen, which is
CC used in the exemplification of the present invention.

XX Sequence 272 AA;

Query Match 100.0%; Score 1397; DB 8; Length 272;
Best Local Similarity 100.0%; Pred. No. 1.1e-136;
Matches 272; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSGGEPMTKTIIVGIDSHAAITAAIMGVDBAISRAVPLVLVSYIKTTHSPDDYRDL 60
DB 1 MSGGEPMTKTIIVGIDSHAAITAAIMGVDBAISRAVPLVLVSYIKTTHSPDDYRDL 60
QY 61 AAHRSLSREASAVBAAGLVKVIETDIPRGAPVLYEASRDAEMICVSGVIGRYASSI 120
DB 61 AAHRSLSREASAVBAAGLVKVIETDIPRGAPVLYEASRDAEMICVSGVIGRYASSI 120
QY 121 LGSTATLEAKACPVAVMRSKVDQPSADINMIVVMTDAPDNEAVLEYAAREAKLRQAP 180
DB 121 LGSTATLEAKACPVAVMRSKVDQPSADINMIVVMTDAPDNEAVLEYAAREAKLRQAP 180
QY 181 ILAAGRREELREIPDGEFERRVQDMHRRHPDVRYPTTHTGTARFLADDERVOLAVI 240
DB 181 ILAAGRREELREIPDGEFERRVQDMHRRHPDVRYPTTHTGTARFLADDERVOLAVI 240
QY 241 GGGAAGOLARLVGSPGHPFRHAECVLYVVR 272
DB 241 GGGAAGOLARLVGSPGHPFRHAECVLYVVR 272

RESULT 2

AD137300
ID AD137300 standard; protein; 279 AA.

AC AD137300;

DT 22-APR-2004 (first entry)

DE M. tuberculosis low oxygen induced antigen Rv2028c SEQ ID NO:21.

KM mycobacterial infection; vaccine; tuberculosis;

KM Mycobacterium tuberculosis; immunisation; antibacterial; gene therapy;

XX low oxygen induced antigen.

XX Mycobacterium tuberculosis.

XX WO2004006952-A2.

XX 22-JAN-2004.

XX 08-JUL-2003; 2003WO-DK000477.

XX 13-JUL-2002; 2002DK-00001098.

XX (STAT-) STATENS SERUM INST.

XX Andersen P, Rosenkrands I, Stryhn A;

XX WPI; 2004-122778/12.

XX DR N-PSDB; AD137345.

PT Use of one or more polypeptides or their fragments, which are expressed
PT during the latent stage of the mycobacterial infection, and/or nucleic
PT acids encoding the polypeptides, for a therapeutic vaccine against

PT tuberculosis.

XX Claim 3; SEQ ID NO 21; 76p; English.

XX The present invention describes polypeptides or their fragments, which
XX are expressed during the latent stage of a mycobacterial infection,
XX and/or nucleic acids encoding the polypeptides, which are useful for
XX creating a therapeutic vaccine against tuberculosis. Also described: (1)
XX a therapeutic vaccine against tuberculosis comprising one or more
XX polypeptides; (2) a method for treating an animal, including a human
XX being, with tuberculosis caused by virulent mycobacteria, e.g. by
XX Mycobacterium tuberculosis, M. africanum or M. bovis; (3) a method for
XX immunising an animal, including a human being, against tuberculosis
XX caused by virulent mycobacteria; (4) a method of diagnosing tuberculosis
XX caused by virulent mycobacteria in an animal, including a human being;
XX (5) a method for diagnosing previous or ongoing infection with a virulent
XX mycobacterium; and (6) a method of diagnosing Mycobacterium tuberculosis
XX infection in a subject. The polypeptides have antibacterial activities,
XX and can be used in vaccines and in gene therapy. The polypeptides are
XX useful for the manufacture of a therapeutic vaccine for treating an
XX individual who is infected by a virulent mycobacterium, e.g. M.
XX tuberculosis, and who is not vaccinated with BCG against tuberculosis.
XX The present sequence represents a low oxygen induced antigen, which is
XX used in the exemplification of the present invention.

XX Sequence 279 AA;

Query Match 29.8%; Score 416; DB 8; Length 279;
Best Local Similarity 36.8%; Pred. No. 2.2e-34;
Matches 98; Conservative 46; Mismatches 116; Indels 6; Gaps 2;

QY 11 TIIVGIDSHAAITAAIMGVDBAISRAVPLVLVSYIKTTHSPDDYD--RDLAAHRSLS 67
DB 9 SIYVIGIDGSKPAVOALMAVDEAASRDIPRLTYALEPDDPGYAAAGAAARKLAAEMNAV 68
QY 68 REASAVBAAGLVKVIETDIPRGAPVLYEASRDAEMICVSGVIGRYASSITLGSTATE 127
DB 69 RYAFVAEADRPVKVEITQERPVTSILRAASAAALVCGAIVGHHFPERVSGTAA 128
QY 128 LAEKAHCPVAVMRSKVDQPSADINMIVVMTDAPDNEAVLEYAAREAKLRQAPILAGR 187
DB 129 LALSACCPVAVMRSKVDQPSADINMIVVMTDAPDNEAVLEYAAREAKLRQAPILAGR 188
QY 188 PEELREIPD--GEFERRVQDMHRRHPDVRYPTTHTGTARFLADDERVOLAVIGGG 244
DB 189 QSGVGTGDDVRLARVQRPYDVVQAAVHGEILDYLAGRGVHVVVLSASD 248
QY 245 AGOLARLVGSPGHPFRHAECVLYV 270
DB 249 QEHVEQLVGAPGNAVLQEAAGCTLLV 274

RESULT 3

AD137321
ID AD137321 standard; protein; 268 AA.

AC AD137321;

DT 22-APR-2004 (first entry)

DE M. tuberculosis low oxygen induced antigen Rv3134c SEQ ID NO:42.

KM mycobacterial infection; vaccine; tuberculosis;

KM Mycobacterium tuberculosis; immunisation; antibacterial; gene therapy;

XX low oxygen induced antigen.

XX Mycobacterium tuberculosis.

XX WO2004006952-A2.

XX 22-JAN-2004.

XX 08-JUL-2003; 2003WO-DK000477.

XX 13-JUL-2002; 2002DK-00001098.
PR (STAT-) STATENS SERUM INST.
PA
XX Andersen P, Rosenkrands I, Stryhn A;
XX WPI; 2004-122778/12.
DR N-PSDB; ADI37366.
XX
PT Use of one or more polypeptides or their fragments, which are expressed
PT during the latent stage of the mycobacterial infection, and/or nucleic
PT acids encoding the polypeptides, for a therapeutic vaccine against
PT tuberculosis.
XX
PS Claim 3; SEQ ID NO 42; 76pp; English.
XX
CC The present invention describes polypeptides or their fragments, which
CC are expressed during the latent stage of a mycobacterial infection,
CC and/or nucleic acids encoding the polypeptides, which are useful for
CC creating a therapeutic vaccine against tuberculosis. Also described: (1)
CC a therapeutic vaccine against tuberculosis comprising one or more
CC polypeptides; (2) a method for treating an animal, including a human
CC being, with tuberculosis caused by virulent mycobacteria, e.g. by
CC Mycobacterium tuberculosis, M. africanum or M. bovis; (3) a method for
CC immunising an animal, including a human being, against tuberculosis
CC caused by virulent mycobacteria; (4) a method of diagnosing tuberculosis
CC caused by virulent mycobacteria in an animal, including a human being;
CC (5) a method for diagnosing previous or ongoing infection with a virulent
CC mycobacterium; and (6) a method of diagnosing Mycobacterium tuberculosis
CC infection in a subject. The polypeptides have antibacterial activities,
CC and can be used in vaccines and in gene therapy. The polypeptides are
CC useful for the manufacture of a therapeutic vaccine for treating an
CC individual who is infected by a virulent mycobacterium, e.g. M.
CC tuberculosis, and who is not vaccinated with BCG against tuberculosis.
CC The present sequence represents a low oxygen induced antigen, which is
CC used in the exemplification of the present invention.
XX
SQ Sequence 268 AA;
Query Match 25.6%; Score 357.5; DB 8; Length 268;
Best Local Similarity 36.5%; Pred. No. 2.7e-28;
Matches 100; Conservative 35; Mismatches 110; Indels 29; Gaps 6;
QY 10 KTIIVGDSHAATTAALMGVDEAISRRAVPLRVSVIKTPHSPDDVDRDLAHSRRE 69
DB 8 RAVVVGIGSRPAATLAALMAVDEAVNRDIPRLVTVIDPSQLSAAEGGCGSARAAALHD 67
QY 70 AQSVAEAGKLVKLTETDIPRGAPGVLVEASRDAMICVSGVIGRVAASIIIGSTATELA 129
DB 68 ASRYVEATGQPVKLTETVLCGRPLTKMQSRSAAMLCVSGVGLD-HYRGRRGSVAATLA 126
QY 130 EKAHCPVAVMSKVDQPA--SDIMKIVVRMTDAPNEAVLEAAREALRQAPILALGGR 187
DB 127 GSALCPVAVIHPSRPAETTSQVSAVVAEV---DNGVLRHAPEEARLRGVPLAAVA-- 180
QY 188 PEELREITPDG-----EPEERVDMMHRRPVRVYVPTTTHGIALRPLADDERVQL 237
DB 181 -VHALETDDVDEQSGRLAHVHSRLAHMTLRLYFEVRDRALAGGSACRHLAANKPQL 239
QY 238 AVIGGEGAGOLARLVGPSGHVFRHAEGSVLVVR 271
DB 240 FVADSHSAHELCAVQPG-----CAVLTVR 264
RESULT 4
ADI37298
ID ADI37298 standard; protein; 295 AA.
XX
XX ADI37298;
AC
XX
XX 22-APR-2004 (first entry)
DT
XX

DE M. tuberculosis low oxygen induced antigen Rv2050c SEQ ID NO:19.
XX mycobacterial infection; vaccine; tuberculosis;
XX Mycobacterium tuberculosis; immunisation; antibacterial; gene therapy;
KW low oxygen induced antigen.
XX
OS Mycobacterium tuberculosis.
XX
XX WQ2004006952-A2.
XX
XX 22-JAN-2004.
XX
XX 08-JUL-2003; 2003WO-DK00477.
XX
XX 13-JUL-2002; 2002DK-00001098.
XX
XX (STAT-) STATENS SERUM INST.
XX
XX Andersen P, Rosenkrands I, Stryhn A;
XX WPI; 2004-122778/12.
DR N-PSDB; ADI37343.
XX
PT Use of one or more polypeptides or their fragments, which are expressed
PT during the latent stage of the mycobacterial infection, and/or nucleic
PT acids encoding the polypeptides, for a therapeutic vaccine against
PT tuberculosis.
XX
PS Claim 3; SEQ ID NO 19; 76pp; English.
XX
CC The present invention describes polypeptides or their fragments, which
CC are expressed during the latent stage of a mycobacterial infection,
CC and/or nucleic acids encoding the polypeptides, which are useful for
CC creating a therapeutic vaccine against tuberculosis. Also described: (1)
CC a therapeutic vaccine against tuberculosis comprising one or more
CC polypeptides; (2) a method for treating an animal, including a human
CC being, with tuberculosis caused by virulent mycobacteria, e.g. by
CC Mycobacterium tuberculosis, M. africanum or M. bovis; (3) a method for
CC immunising an animal, including a human being, against tuberculosis
CC caused by virulent mycobacteria; (4) a method of diagnosing tuberculosis
CC caused by virulent mycobacteria in an animal, including a human being;
CC (5) a method for diagnosing previous or ongoing infection with a virulent
CC mycobacterium; and (6) a method of diagnosing Mycobacterium tuberculosis
CC infection in a subject. The polypeptides have antibacterial activities,
CC and can be used in vaccines and in gene therapy. The polypeptides are
CC useful for the manufacture of a therapeutic vaccine for treating an
CC individual who is infected by a virulent mycobacterium, e.g. M.
CC tuberculosis, and who is not vaccinated with BCG against tuberculosis.
CC The present sequence represents a low oxygen induced antigen, which is
CC used in the exemplification of the present invention.
XX
SQ Sequence 295 AA;
Query Match 20.5%; Score 286; DB 8; Length 295;
Best Local Similarity 29.6%; Pred. No. 8.7e-21;
Matches 85; Conservative 48; Mismatches 126; Indels 28; Gaps 8;
QY 12 IIVGIDGSHAATTAALMGVDEAISRRAVPLRVSVIK-----PTHPSPDDV-----DRDL 60
DB 10 VVGVDSLSLSDAABACGATDPAAMRNIPLYVHVAVNDVATWPMPTPEITMGVQDEBGR 69
QY 61 AHAERSLRQAQSAVEAAGKLVKLTETDIPRGAPGVLVEASRDAMICVSGVIGRVAASI 120
DB 70 QIVANNAVLAKEAAGADRKI-SVXSELVFSTPVPVTWEISNEAEVVLGSSGAGALARGL 128
QY 121 LGSTATLEAKAHCPVAVMSKVDQPASDINW--IIVRMTDAPNEAVLEAAREALRQ 178
DB 129 LGSVSSSLVRRACGPVAVIHSD--DAVLPDQAHAVLIGIGSPVSELATAVAFDEARRG 187
QY 179 APIALGGRPE-ELREIPDGEFB-----RRVQDMHRRPVRVYVPTTTHGIALR 226
DB 188 VELTAIVHMSDVEVVELPGIDFSAVQDEALSLAERLAGNQERIPDPVSRVVVCDRPAR 247

QY 227 FLADHDERVOLAVTGG-GEAGQLARLVGSGHPVRHAECSVLVYR 272
 DB 248 KLVGKASASQQLVVGSHGRGGTLTKMLGSSNAVLHAAVPIYARQ 294

RESULT 5
 ID AAY87878 standard; protein; 297 AA.
 AC AAY87878;
 DT 06-OCT-2000 (first entry)
 DE M. tuberculosis antigen TB32 protein.
 XX
 XX Tuberculosis; TB; antigen; vaccine; diagnosis; somatic; tuberculostatic;
 KM infection; interferon-gamma; IFN-gamma; protective immunity; therapy;
 KM delayed type hypersensitivity response; TB32.
 XX
 OS Mycobacterium tuberculosis.
 XX
 PN MO200021983-A2.
 XX
 PD 20-APR-2000.
 XX
 PP 08-OCT-1999; 99MO-DK000538.
 XX
 PR 08-OCT-1998; 98DK-00001281.
 PR 21-JAN-1999; 99US-011673P.
 XX
 PA (STAT-) STRATENS SERUM INST.
 XX
 PI Andersen P, Weidling K, Hansen CV, Florio W, Okkels LM;
 PI Skjot RLV, Rosenkrands I;
 XX
 DR WPI: 2000-317931/27.
 DR N-PSDB: AAA39572.
 XX
 PT Novel polypeptide of somatic protein extract useful as vaccine against
 PT virulent Mycobacterium infection, isolated from cell wall, cell membrane
 PT and cytosol.
 PS
 PS Claim 1; Page 106-107; 126pp; English.

CC This invention describes a novel polypeptide (PP) of somatic proteins
 CC extract (I) which have tuberculostatic activity. (I) or their subsequence
 CC has at least one of the following properties: (a) the PP induces an in
 CC vitro recall response, or an in vitro response, during primary infection
 CC with virulent Mycobacterium, determined by a release of interferon (IFN)-
 CC gamma, (b) PP induces a protective immunity, determined by vaccinating an
 CC animal with PP and an adjuvant, three times at two weeks intervals, (c)
 CC PP induces an in vitro response, or in vitro recall response, determined
 CC by release of IFN-gamma of at least 1000 pg/ml or 500 pg/ml,
 CC respectively, from Peripheral Blood Mononuclear Cells (PBMC) withdrawn
 CC from TB patients, or PPB positive individuals, 6 months after diagnosis,
 CC (d) PP induces a specific antibody response in a TB patient, as
 CC determined by enzyme linked immunosorbent assay (ELISA) technique or a
 CC western blot, (e) PP induces a positive delayed type hypersensitivity
 CC (DTH) response, determined by intradermal injection. (I) and (II) are
 CC useful in preparing a prophylactic or therapeutic medicine as a vaccine
 CC for induction of a protective or generation of an immune response in a
 CC mammal against infection with a virulent Mycobacterium. (I) and (II) are
 CC also useful as diagnostic reagent for the diagnosis of a virulent
 CC Mycobacterium infection. The vaccine of the invention induces efficient
 CC immunological memory, providing long term protection against TB. This
 CC sequence represents a Microbacterium tuberculosis TB32 antigen described
 CC in the invention
 XX
 SQ Sequence 297 AA;

Query Match 19.1%; Score 267; DB 3; Length 297;
 Best Local Similarity 29.3%; Pred. No. 8.4e-19;
 Matches 84; Conservative 42; Mismatches 131; Indels 30; Gaps 8;

QY 12 IYVIGDSHAALITAAWGVDEDAISRVPRLVSVYK-----THSEP-----DDYDR 58
 DB 10 IIVGIDSPAAQAVAVRWAAADAEIRKIPLTILVAHSVPEVATWLEVPPLPQVLRMOQDHR 69

QY 59 DLHAERSLREASVAEAAAGKLVKIENDIPRGAPVLYEASDAEMICVSGVIGRYAS 118
 DB 70 HL--IDDLKLVVEQASLRAGP-PTVHSEIVPAAAVPTLVMSKDAVLMTGVGCGSGRMPG 126

QY 119 SILGSTATBELAKRACVAVMRSKVD-QPASDINMIVVRMTDAPDNEAVLYEAREAKLR 177
 DB 127 RLIGSVSSGLLRHACHCVVLIHDEDSVMPHQAPVLYGVGDSASSELATATAIFDEASRR 186

QY 178 QAPITALGRPE-ELKEIP-----DGEFERRVQDMHHRPDPVRYPTTHTGIA 225
 DB 187 NVPLVLAHMSVDVSEMPGIDWPATQSMAEQVLARLGMQSRYPVATRRVAVDQPA 246

QY 226 RFLADHDERVOLAVTGG-GEAGQLARLVGSGHPVRHAECSVLVYR 271
 DB 247 KQVQRSEBAQLVVGSRGRGYAGMLVGSGETVAQLARTPIYAR 293

RESULT 6
 ID ABG30907 standard; protein; 297 AA.
 AC ABG30907;
 DT 21-OCT-2002 (first entry)
 DE Mycobacterium tuberculosis dormancy-induced protein Rv2623.
 XX
 XX Anti-mycobacterial; Mycobacterium tuberculosis; strain H37RV;
 KM non-oxygen limiting; hypoxic stationary; hypoxic growth phase;
 KM tuberculosis; dormant mycobacterial infection; dormancy-induced; Rv2623;
 KM tuberculostatic; vaccine.
 XX
 OS Mycobacterium tuberculosis.
 XX
 PN MO200248391-A2.
 XX
 PD 20-JUN-2002.
 XX
 PP 11-DEC-2001; 2001MO-BE014551.
 XX
 PR 13-DEC-2000; 2000GB-00030368.
 XX
 PA (MOLE-) INST MOLECULAR & CELL BIOLOGY.
 XX
 PI Dick T, Calvin BKK;
 XX
 DR WPI: 2002-583530/62.
 DR N-PSDB: ABK89546.
 XX
 PT Identifying an anti-mycobacterial agent that modulates
 PT activity/expression of a protein expressed by Mycobacterium, involves
 PT protein or polynucleotide/vector encoding it.
 PS
 PS Example 2; Page 44-45; 56pp; English.

CC The present invention relates to a new method for the identification of
 CC anti-mycobacterial agent that modulates activity and/or expression of
 CC protein expressed by Mycobacterium in non-oxygen limiting or hypoxic
 CC stationary, hypoxic growth phase. The method involves contacting a test
 CC agent and protein, and monitoring the effect of agent on activity/
 CC expression of the protein. The methods of the invention are useful for
 CC treating a human or animal body by therapy, in a diagnostic method
 CC practised on the human or animal body, and for manufacturing medicament
 CC for diagnosis, prophylaxis or treatment of mycobacterial infection,
 CC especially tuberculosis. The invention is also useful for in vitro or in
 CC vivo diagnosing of mycobacterial infection in a sample. The invention can
 CC be used for diagnosing a dormant mycobacterial infection. The present

CC amino acid sequence represents the Mycobacterium tuberculosis strain
 CC H37Rv dormancy-induced RV2623 protein of the invention
 XX
 SQ Sequence 297 AA;

Query Match 19.1%; Score 267; DB 5; Length 297;
 Best Local Similarity 29.3%; Pred. No. 8.4e-19;
 Matches 84; Conservative 42; Mismatches 131; Indels 30; Gaps 8;

QY IIVGIDGSHAAITAAALMGVDEAISRAPLRLVSVIKP-----THPSP-----DDYDR 58
 DB IIVGIDGSPAAQAVAVRAARDAHELRKIPLTLVHNVSEVAATWLEVPGLRWQODHGR 69
 QY DLAAERSLREASQSAVEAAGKLVKIETDIPRGAPGVLEASRDAMI CVSGVIGRYAS 118
 DB 70 HL--IDDLAKVVEQASLRAGP-PTVHSEIVPAAAVPTLVDSKDAVLVVGGLSGGRMPG 126
 QY SIIGSTATELAEKACPVAVWRKSYD-QPAGDINIVVRMDAPNEAVLEAAREAKLR 177
 DB 127 RLTSVSSGGLRHAHCPVVIIDHEDSVMPHPQAPVLGVGSSASELATATAFDEASRR 186
 QY 178 QAPILALGGRPE-ELREIP-----DGEFERRVQDMHRRHDPVVRVPIITHTGIA 225
 DB 187 NVDLVLAHMSDVDSVSEWPGCIDMPATQSMAEQVLAERLAGMOERYENVAITRVVVRDQPA 246
 QY 226 RFLADHDERVOLAVIGG-GEAGQLARLVGPSGHPVFRHAEGSVLVVR 271
 DB 247 RQIVGRSEBAQLVNVGSRGRGVAGMLVGSVGETVAQALARPVIYAR 293

RESULT 7
 AAB79640
 ID AAB79640 standard; protein; 301 AA.

AC AAB79640;

XX 30-APR-2001 (first entry)

DE Corynebacterium glutamicum MP protein sequence SEQ ID NO:14.

XX
 KW Corynebacterium glutamicum; metabolic pathway protein; MP protein;
 KW fine chemical production; microorganism; organic acid; nucleoside;
 KW nonproteinogenic amino acid; purine base; pyrimidine base; nucleotide;
 KW lipid; saturated fatty acid; unsaturated fatty acid; diol; vitamin;
 KW carbohydrate; aromatic compound; cofactor; polypeptide; enzyme.

OS Corynebacterium glutamicum.

XX WO200100843-A2.

XX 04-JAN-2001.

XX 23-JUN-2000; 2000WO-1B000923.

XX 25-JUN-1999; 99US-0141031P.
 PR 01-JUL-1999; 99US-01030476.
 PR 02-JUL-1999; 99US-0142101P.
 PR 08-JUL-1999; 99DE-01031415.
 PR 08-JUL-1999; 99DE-01031418.
 PR 08-JUL-1999; 99DE-01031419.
 PR 08-JUL-1999; 99DE-01031420.
 PR 08-JUL-1999; 99DE-01031424.
 PR 08-JUL-1999; 99DE-01031428.
 PR 08-JUL-1999; 99DE-01031434.
 PR 08-JUL-1999; 99DE-01031435.
 PR 08-JUL-1999; 99DE-01031443.
 PR 08-JUL-1999; 99DE-01031453.
 PR 08-JUL-1999; 99DE-01031457.
 PR 08-JUL-1999; 99DE-01031465.
 PR 08-JUL-1999; 99DE-01031478.
 PR 08-JUL-1999; 99DE-01031510.
 PR 08-JUL-1999; 99DE-01031541.
 PR 08-JUL-1999; 99DE-01031573.

PR 08-JUL-1999; 99DE-01031592.
 PR 08-JUL-1999; 99DE-01031632.
 PR 08-JUL-1999; 99DE-01031634.
 PR 08-JUL-1999; 99DE-01031636.
 PR 09-JUL-1999; 99DE-01032125.
 PR 09-JUL-1999; 99DE-01032126.
 PR 09-JUL-1999; 99DE-01032130.
 PR 09-JUL-1999; 99DE-01032186.
 PR 09-JUL-1999; 99DE-01032206.
 PR 09-JUL-1999; 99DE-01032227.
 PR 09-JUL-1999; 99DE-01032228.
 PR 09-JUL-1999; 99DE-01032229.
 PR 09-JUL-1999; 99DE-01032230.
 PR 09-JUL-1999; 99DE-01032392.
 PR 14-JUL-1999; 99DE-01032926.
 PR 14-JUL-1999; 99DE-01032928.
 PR 14-JUL-1999; 99DE-01033004.
 PR 14-JUL-1999; 99DE-01033005.
 PR 14-JUL-1999; 99DE-01033006.
 PR 14-JUL-1999; 99US-0148613P.
 PR 12-AUG-1999; 99DE-01040764.
 PR 27-AUG-1999; 99DE-01040765.
 PR 27-AUG-1999; 99DE-01040766.
 PR 27-AUG-1999; 99DE-01040832.
 PR 31-AUG-1999; 99DE-01041378.
 PR 31-AUG-1999; 99DE-01041379.
 PR 31-AUG-1999; 99DE-01041380.
 PR 31-AUG-1999; 99DE-01041394.
 PR 31-AUG-1999; 99DE-01041396.
 PR 03-SEP-1999; 99DE-01042076.
 PR 03-SEP-1999; 99DE-01042077.
 PR 03-SEP-1999; 99DE-01042079.
 PR 03-SEP-1999; 99DE-01042086.
 PR 03-SEP-1999; 99DE-01042087.
 PR 03-SEP-1999; 99DE-01042088.
 PR 03-SEP-1999; 99DE-01042095.
 PR 03-SEP-1999; 99DE-01042124.
 PR 03-SEP-1999; 99DE-01042129.
 PR 09-MAR-2000; 2000US-0187970P.

XX (BADI) BASF AG.

XX Pompejus M, Kroegeer B, Schroeder H, Zelder O, Haberhauer G;

XX WPI; 2001-137957/14.

DR N-PSDB; AAF71759.

XX
 PT Nucleic acids from Corynebacterium glutamicum encoding metabolic pathway
 PT proteins, useful for producing fine chemicals in microorganisms,
 PT including organic acids, nonproteinogenic amino acids, and purine and
 PT pyrimidine bases.

PS Claim 20; Page 174-175; 1737pp; English.

XX AAF71753 to AAF72330 encode the Corynebacterium glutamicum metabolic
 CC pathway (MP) proteins given in AAB79634 to AAB80211. The C. glutamicum MP
 CC nucleic acids are useful for the production of fine chemicals in
 CC microorganisms, including organic acids, nonproteinogenic amino acids,
 CC purine and pyrimidine bases, nucleosides, nucleotides, lipids, saturated
 CC and unsaturated fatty acids, diols, carbohydrates, aromatic compounds,
 CC vitamins, cofactors, polypeptides and enzymes

XX Sequence 301 AA;

Query Match 19.0%; Score 266; DB 4; Length 301;
 Best Local Similarity 28.8%; Pred. No. 1.1e-18;
 Matches 85; Conservative 45; Mismatches 131; Indels 34; Gaps 8;

QY 8 TMKTIIVGIDGSHAAITAAALMGVDEAISRAPLRLVSVIKPTSPS-----PDDYDR 58
 DB 3 TEDIVVAVDSDSKQAVRAARDAHELRKIPLTLVHNVSEVAATWLEVPGLRWQODHGR 60
 QY 59 DL-AAERSLREASQSAVEAAGKLVKIETDIPRGAPGVLEASRDAMI CVSGVIGRYA 117

```

Db      61 DLQEALEKINEADIDAEVAPETIKGHTIIEGSDIDMLEMSPDATIIVWGSSGLGELS 120
Oy      118 SSISGSTATBLAEKACPVAVMR--SKVDPASDINMIIVMTDAPDNEAVLEAAREAK 175
Db      121 GMMWGSVGAIVSHAKCPVVVVRSDSAVNEP-SKGPVVVGVDGSEVSGQATEYAFAEAE 179
Oy      176 LR-----QAPILALGGRPEBELREIPDGEFF--RRVQDMHRRHPDVRYP 217
Db      180 ARGAEIVAVHTMDMVOVASLAGLAAAOQMDVERQOTDMLIERLAPLVEKYPSTVVK 239
Oy      218 ITHHTGIARFLADDERVQLAVIGG-GEAGQLARLVGSGHPVRHAECSVLVYR 271
Db      240 IITRDRPVRLALAEASENAOQLLVGSHGRGFKGMILGSTSRALLQSAFCPMVYR 294

RESULT 8
AAU71872
ID      AAU71872 standard; protein; 301 AA.
XX
AC      AAU71872;
XX
DT      26-FEB-2002 (first entry)
XX
DE      C. glutamicum metabolic pathway protein encoded by gene #7.
XX
KW      Metabolic pathway protein; MP; lysine biosynthesis pathway;
KW      methionine biosynthesis pathway; large-scale production of fine chemical;
KW      Corynebacterium diphtheriae; diphtheria.
XX
OS      Corynebacterium glutamicum.
XX
PN      MO200166573-A2.
XX
PD      13-SEP-2001.
XX
PF      22-DEC-2000; 2000MO-IB002035.
XX
PR      09-MAR-2000; 2000US-0187970P.
XX      23-JUN-2000; 2000US-00606740.
XX
PA      (BAD1 ) BASF AG.
XX
PI      Pompejus M, Kroeger B, Schroeder H, Zelder O, Haberhauer G;
XX      Kim J, Lee H, Hwang B;
XX
DR      WPI; 2001-582269/65.
XX      N-PSDB; AAS96082.
XX
PT      Nucleic acids encoding metabolic pathway proteins from Corynebacterium
PT      glutamicum, useful for producing methionine and lysine in Corynebacterium
PT      and Brevibacterium.
XX
PS      Disclousure; Page 169-170; 316pp; English.
XX
CC      The present invention relates to the isolation of novel Corynebacterium
CC      glutamicum genes (AAS96073-AAS96132) encoding metabolic pathway (MP)
CC      proteins. The metabolic pathway proteins of the invention include enzymes
CC      involved in the lysine and methionine biosynthetic pathways. The
CC      polynucleotide sequences of the invention can be used for the large-scale
CC      production and/or modulation of expression of fine chemicals such as
CC      lysine and methionine. The sequences of the invention may be used to
CC      identify C. glutamicum and related organisms e.g. C. diphtheriae in a
CC      subject to detect diphtheria. AAU71863-AAU71922 represent the novel C.
CC      glutamicum metabolic pathway proteins of the invention
XX
XX
SQ      Sequence 301 AA;
Oy      Query Match      19.0%; Score 266; DB 4; Length 301;
Oy      Best Local Similarity 28.8%; Pred. No. 1,1e-18;
Oy      Matches 85; Conservative 45; Mismatches 111; Indels 34; Gaps 8;
Oy      8 TMKTIIVGIDGSHAAITPAALMGVDEAISRVAVPLRLVSVIKPTPS-----PDDYDR 58

```

```

Db      3 TEDIVVAVVDGSPASAKOAVRMAANTANKRGIPRLALS--SYTFPOFLYAGWVPPELFD 60
Oy      59 DL-AHARSRLREAOASAVEAGKLVKLETDI-PRRPAGPVLEASRDMEICVSGVIGRYA 117
Db      61 DLQEALEKINEADIDAEVAPETIKGHTIIEGSDIDMLEMSPDATIIVWGSSGLGELS 120
Oy      118 SSISGSTATBLAEKACPVAVMR--SKVDPASDINMIIVMTDAPDNEAVLEAAREAK 175
Db      121 GMMWGSVGAIVSHAKCPVVVVRSDSAVNEP-SKGPVVVGVDGSEVSGQATEYAFAEAE 179
Oy      176 LR-----QAPILALGGRPEBELREIPDGEFF--RRVQDMHRRHPDVRYP 217
Db      180 ARGAEIVAVHTMDMVOVASLAGLAAAOQMDVERQOTDMLIERLAPLVEKYPSTVVK 239
Oy      218 ITHHTGIARFLADDERVQLAVIGG-GEAGQLARLVGSGHPVRHAECSVLVYR 271
Db      240 IITRDRPVRLALAEASENAOQLLVGSHGRGFKGMILGSTSRALLQSAFCPMVYR 294

RESULT 9
AAU71865
ID      AAU71865 standard; protein; 301 AA.
XX
AC      AAU71865;
XX
DT      26-FEB-2002 (first entry)
XX
DE      C. glutamicum metabolic pathway protein encoded by RXA00657 gene.
XX
KW      Metabolic pathway protein; MP; lysine biosynthesis pathway;
KW      methionine biosynthesis pathway; large-scale production of fine chemical;
KW      Corynebacterium diphtheriae; diphtheria.
XX
OS      Corynebacterium glutamicum.
XX
PN      MO200166573-A2.
XX
PD      13-SEP-2001.
XX
PF      22-DEC-2000; 2000MO-IB002035.
XX
PR      09-MAR-2000; 2000US-0187970P.
XX      23-JUN-2000; 2000US-00606740.
XX
PA      (BAD1 ) BASF AG.
XX
PI      Pompejus M, Kroeger B, Schroeder H, Zelder O, Haberhauer G;
XX      Kim J, Lee H, Hwang B;
XX
DR      WPI; 2001-582269/65.
XX      N-PSDB; AAS96075.
XX
PT      Nucleic acids encoding metabolic pathway proteins from Corynebacterium
PT      glutamicum, useful for producing methionine and lysine in Corynebacterium
PT      and Brevibacterium.
XX
PS      Claim 24; Page 150-151; 316pp; English.
XX
CC      The present invention relates to the isolation of novel Corynebacterium
CC      glutamicum genes (AAS96073-AAS96132) encoding metabolic pathway (MP)
CC      proteins. The metabolic pathway proteins of the invention include enzymes
CC      involved in the lysine and methionine biosynthetic pathways. The
CC      polynucleotide sequences of the invention can be used for the large-scale
CC      production and/or modulation of expression of fine chemicals such as
CC      lysine and methionine. The sequences of the invention may be used to
CC      identify C. glutamicum and related organisms e.g. C. diphtheriae in a
CC      subject to detect diphtheria. AAU71863-AAU71922 represent the novel C.
CC      glutamicum metabolic pathway proteins of the invention
XX
XX
SQ      Sequence 301 AA;
Oy      Query Match      19.0%; Score 266; DB 4; Length 301;

```

Best Local Similarity 28.8%; Pred. No. 1.1e-18;
Matches 85; Conservative 45; Mismatches 131; Indels 34; Gaps 8;

```

Qy      8  MMKTIIVGDSHAITLALMGVDEIASVAPRLTYSVILKPTHS-----PDDYR  58
Db      3  TEDIIVAVDGDASQAVRMAANTANKGIFRLMS--SYMPQFLYAEGVPPQSLFD  60
Qy      59  DL-AHAERSLREAGSAVEAAGKLVKLTETDIPRPGAPVLVEASRDAEMTCVSGVIGRYA  117
Db      61  DLQAEALKEINERDIAHEVAPEIKGHTIAEGSPIDMLLEMSPDATMIWMSRGIGLS  120
Qy      118  SSILGSTATELAEKXHCPLVAVNR--SKYDQPSADIMVIYVRMTADAQNEAVLEIYAREAK  175
Db      121  GMWMSVSGAVVSHAKCPVVVVRBDSAVED--SKYSPVVAVGVDSSESOQATYAAREAE  179
Qy      176  LR-----CAPILALGGRPEELREIPDGEFE--RAYQDMHHRPDRVVP  217
Db      180  ARGAEIVAVHTMMDMQVQASLGLAAAOQOMPEVEHQQIDMLIERLALPVEKIPSTLYVK  239
Qy      218  ITTHTGIAFLADHDREVOLAVIGC-GEAGQOLARLVGSGHPVFRIAEGSVLYVR  271
Db      240  IITRDPRVALAEASENAGLLVVGSHRGGFKMLIGTSRRLLQAGPCPMVVR  294

```

RESULT 10
AAG92991
ID AAG92991 standard; protein; 301 AA.

AC AAG92991;

DT 26-SEP-2001 (first entry)

DE C glutamicum protein fragment SEQ ID NO: 6745.

KW Coryneform bacterium; amino acid synthesis; vitamin; saccharide; organic acid synthesis.

Corynebacterium glutamicum.

PN EP1108790-A2.

PD 20-JUN-2001

PF 18-DEC-2000; 2000EP-00127688.

PR 16-DEC-1999; 99JP-00377484.

PR 03-AUG-2000; 2000JP-00280988

PA (KYOW) KYOWA HAKKO KOGYO KK

PI Nakagawa S, Mizoguchi H, Akiyama T, et al. 2000. The

XX

DR N-PSDB; AAH68210.

Novel polynucleotides derived

PT expression profile or pattern

PS Claim 17; SEQ ID NO 6745; 246pp + Sequence Listing; English.

CC The present invention provides a number of nucleotide and protein
CC sequences from the *Corynebacterium* *Corynebacterium glutamicum*. These
CC are useful for identifying the mutation point of a gene derived from a
CC mutant of *Corynebacterium* *Corynebacterium*, measuring expression amount and analyzing
CC the expression profile or expression pattern of a gene derived from
CC *Corynebacterium* *Corynebacterium*, and identifying a homologue of a gene derived from
CC *Corynebacterium* *Corynebacterium*. *Corynebacterium* bacteria are useful for producing amino
CC acids, nucleic acids, vitamins, saccharides and organic acids,
CC particularly L-lysine. The present sequence is a protein described in the
CC exemplification of the invention. Note: The sequence data for this patent

CC did not form part of the printed specification, but was obtained in
CC electronic format directly from the European Patent Office
XX
50 Sequence 301 AA;

Query Match	19.0%	Score 266	DB 4	Length 301
-------------	-------	-----------	------	------------

Matches 85; Conservative 45; Mismatches 131; Indels 34; Gaps 8;

```

Qy      8  TMTKIIIVGIDSHAAITLALMGVDBALSRAVLELVSVIKRTHS-----PDDYDR  58
           :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      3  TEDIIVAAVDDSGDASQAVRMAMANTANKRGILPLKIAS--SYTTFQFIYAGSMYPRDELF  60
           :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

Qy      59  DL-AHAERSLREASQAVENAAKGLVYKIEITDIPIRGAPRYLYEASHDAEMI CVGSGVIGRYA  117
           :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      61  DLQAVALEKINEARDIAHVAPEIKIGHTIAEGSPIDMLTLEMSPDATMI VMGSRGIGGLS  120
           :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

Qy      118  SLLSGTATELLEKAKHCPVAVNR--SKYDQASADINIIIVYRMTPDAPNEAVLEAAREAK  175
           :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      121  GMMGSGVSAVNVSHAKCPVVVVREDASVNEP--SKYGPVVVGDSSEVSQATEAFAFAAE  179
           :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

Qy      176  LR-----QAPILALGGRPEELREIIPDGSFE---RYVDIMNNHNDYRVVYV  217
           :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      180  ARGAEILVAHTIMDMQVQASLAGLAQAQQMDVEVRQQTDLILRLRLVLEKRPSTVYKK  239
           :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

Qy      218  ITTHTGIAPFLADHDVQVLAIVG-GEAGQLARLVESGSHVPRVNAECSVLVYR  271
           :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      240  IITRDPRVALAEASENQILVVGSHGSGKGMILSTSRALQSPRCSPMMVVR  294
           :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

```

RESULT 11
ABG80324
ID ABG80324 standard; protein; 301 AA

AC ABG80324;

DT 15-NOV-2002 (first entry)

DE C. glutamicum metabolic pathway (MP) protein #4.

KM Metabolic pathway protein; MP; cell metabolism; amino acid; vitamin;
KM cofactor; nucleotide; nucleoside; creatinase; fine chemical production
KM organic acid; non-proteinogenic amino acid; purine base; carboxylate
KM pyrimidine base; lipid; unsaturated fatty acid; diol; polyketide;
KM aromatic compound; food industry; animal feed; cosmetic industry;
KM pharmaceutical industry; enzyme.

OS *Corynebacterium glutamicum* ATCC 13032.

PN WO200251231-A1.

PD 04-JUL-2002.

PF 22-DEC-2000; 2000WO-EP013143.

PR 22-DEC-2000; 2000WO-EP013143.

PA (BADI) BASF AG.

PI Pompejus M, Kroeger B, Zelder O, Schroeder H,

DR WPI; 2002-643289/69

XX

PT New metabolic pathway genes of *Corynebacterium glutamicum* for producing
PT fine chemicals, e.g. lipids, (un)saturated fatty acids, vitamins,
PT cofactors or enzymes used in food, feed, cosmetics or pharmaceutical
PT industries.

PS Claim 18; Page 107-108; 176pp; English.

CC The present invention relates to the isolation of *Corynebacterium*
CC glutamicum metabolic pathway (MP) proteins, and the polynucleotide

CC sequences encoding them. The MP proteins are enzymes involved in the
CC metabolism of molecules important for the normal functioning of cells
CC (e.g. amino acids, vitamins, cofactors, nucleosides and nucleosides, or
CC trehalose). The polynucleotide sequences encoding the MP proteins are
CC useful for producing fine chemicals, particularly organic acids, non-
CC proteinogenic amino acids, purine and pyrimidine bases, nucleosides,
CC nucleotides, lipids, (un)saturated fatty acids, diols, carbohydrates,
CC aromatic compounds, vitamins, cofactors, polyketides and enzymes. The
CC fine chemicals are useful in the food, animal feed, cosmetic or
CC pharmaceutical industries. ABG80321-ABG80343 represent the C. glutamicum
CC MP proteins of the invention

XX Sequence 301 AA;

SQ Query Match 19.0%; Score 266; DB 5; Length 301;

Best Local Similarity 28.8%; Pred. No. 1.1e-18;

Matches 85; Conservative 45; Mismatches 131; Indels 34; Gaps 8;

QY 8 TMKTIIVGIDGSHAAITTAALMGVDEAISRAVPLRLVSVIKPTHP8-----PDYDR 58

Db 3 TEDIVVAVVADSDSKQAVRWANANTANKRGIPRLAS--SYTMOFLYABGWPPELFD 60

QY 59 DL-AHAERSLRBAQSAVBAAGLVKIERDIPRGAPVLYEASRDABMICVSGVIGRYA 117

Db 61 DLQBALEKINERKDIARIEVAPEIKIGHTIABGSPIDMLIENSPDATTIMVSGRGGLS 120

QY 118 SSIIGSTATBELAEKACPVAVNR--SKVDQASDINMIIVRTDAPDNEAVLEYAAREAK 175

QY 121 GWMVGSVGAIVSHAKCPVVVVRSDAVNED-SKYGPRVVGVDSEVSGQATEYAFPAEAE 179

Db 176 LR-----QAPILALGGRPELRIRPDGEF--RRVQDMHNRHPRVRYR 217

QY 180 ARGAEIYAVHTMMDQVQASLAGLAQAQQQWDEVERQDTMLIRLAPLVKEXYPTVTKK 239

QY 218 ITTHGTIARPLADDERVOLAVIGG-GEAGQLARLVGSPGHVFRHAECSVLVYR 271

Db 240 IITDRPRVALABASENAQLLVGSHGRGFKMGLIGSTRALLQSAFCPRMAYR 294

Db

RESULT 12

ABE15399 standard; protein; 301 AA.

XX ABE15399;

AC 22-SEP-2005 (first entry)

XX C glutamicum metabolic pathway regulatory (MR) protein SeqID314.

XX metacoliem; microorganism detection; microorganism identification;

XX genome; evolution; protein production.

XX Corynebacterium glutamicum.

OS US2005153402-A1.

XX 14-JUL-2005.

PD 06-DEC-2004; 2004US-00006098.

XX 25-JUN-1999; 99US-0141031P.

XX 01-JUL-1999; 99DE-01030476.

XX 01-JUL-1999; 99US-0142690P.

XX 08-JUL-1999; 99DE-01031419.

XX 08-JUL-1999; 99DE-01031420.

XX 09-JUL-1999; 99DE-01032122.

XX 09-JUL-1999; 99DE-01032128.

XX 09-JUL-1999; 99DE-01032134.

XX 09-JUL-1999; 99DE-01032206.

XX 14-JUL-1999; 99DE-01033003.

XX 27-AUG-1999; 99US-0151251P.

XX 31-AUG-1999; 99DE-01041390.

PR 03-SEP-1999; 99DE-01042088.
PR 03-SEP-1999; 99DE-01042124.
PR 23-JUN-2000; 2000US-00602874.

XX (BADI) BASF AG.

PI Pompejus M, Krogger B, Schroder H, Zelder O, Haberhauer G;

XX WPI; 2005-496831/50.

DR N-PSDB; ABE15398.

XX An isolated metabolic pathway regulatory polypeptide from Corynebacterium

PT glutamicum, its portion, or its variant, useful for diagnosing presence

PT or activity of C. diptheriae in subject.

PS Claim 20; SEQ ID NO 314; 65pp; English.

XX This invention relates to novel isolated metabolic pathway regulatory

CC polypeptides from Corynebacterium glutamicum, designated MR proteins, and

CC the DNA sequences which encode them. The invention is useful in

CC identification of C glutamicum and related organisms, mapping of genomes

CC of organisms related to C glutamicum, identification and localization of

CC C glutamicum sequences of interest, evolutionary studies, determination

CC of MR protein regions required for function, modulation of MR protein

CC activity and modulation of cellular production of desired compounds such

CC as a fine chemical. The present sequence is that of a novel isolated

CC metabolic pathway regulatory polypeptide of the invention from

CC Corynebacterium glutamicum. Note: This sequence does not appear in the

CC specification but was obtained in electronic format directly from USPTO.

SQ Sequence 301 AA;

QY Query Match 19.0%; Score 266; DB 9; Length 301;

Best Local Similarity 28.8%; Pred. No. 1.1e-18;

Matches 85; Conservative 45; Mismatches 131; Indels 34; Gaps 8;

QY 8 TMKTIIVGIDGSHAAITTAALMGVDEAISRAVPLRLVSVIKPTHP8-----PDYDR 58

Db 3 TEDIVVAVVADSDSKQAVRWANANTANKRGIPRLAS--SYTMOFLYABGWPPELFD 60

QY 59 DL-AHAERSLRBAQSAVBAAGLVKIERDIPRGAPVLYEASRDABMICVSGVIGRYA 117

Db 61 DLQBALEKINERKDIARIEVAPEIKIGHTIABGSPIDMLIENSPDATTIMVSGRGGLS 120

QY 118 SSIIGSTATBELAEKACPVAVNR--SKVDQASDINMIIVRTDAPDNEAVLEYAAREAK 175

Db 121 GWMVGSVGAIVSHAKCPVVVVRSDAVNED-SKYGPRVVGVDSEVSGQATEYAFPAEAE 179

QY 176 LR-----QAPILALGGRPELRIRPDGEF--RRVQDMHNRHPRVRYR 217

Db 180 ARGAEIYAVHTMMDQVQASLAGLAQAQQQWDEVERQDTMLIRLAPLVKEXYPTVTKK 239

QY 218 ITTHGTIARPLADDERVOLAVIGG-GEAGQLARLVGSPGHVFRHAECSVLVYR 271

Db 240 IITDRPRVALABASENAQLLVGSHGRGFKMGLIGSTRALLQSAFCPRMAYR 294

Db

RESULT 13

ABU25937 standard; protein; 300 AA.

XX ABU25937;

XX 19-JUN-2003 (first entry)

XX Protein encoded by prokaryotic essential gene #11464.

XX Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX Corynebacterium diptheriae.

XX WO200277183-A2.

PD 03-OCT-2002.
 XX 21-MAR-2002; 2002WO-US009107.
 XX 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX Wang L, Zamudio C, Malone C, Hasselbeck R, Ohlsen KL, Zykkind JW,
 P1 Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH,
 P1 WPI: 2003-029926/02.
 DR N-PSDB; ACN29807.
 DR
 PT New antisense nucleic acid, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX
 PS Claim 25; SEQ ID NO 53861; 1766pp; English.
 XX
 CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 CC
 XX Sequence 300 AA;
 SO
 Query Match 17.8%; Score 249; DB 6; Length 300;
 Best Local Similarity 28.5%; Pred. No. 6,4e-17;
 Matches 84; Conservative 46; Mismatches 123; Indels 42; Gaps 10;
 Oy 12 IIVGIDGSHAHTALMGVDEAISRAVPLRLSVIKPHPPDPD-YDRDLHAHRSLSMEA 70
 Db 7 VVVAVDGSEASONAVRAMANTANKGVPRLA---SYTFPQFIYAGCMPPOGLFPEL 62
 Oy 71 QS-----AVEAAGKL-----VKIETDIPRGPAGPVLVEASRDAMICVSGVIGRYASS 119
 Db 63 QSEITMDMEARVVAHEVAPDKIGYVAEESPIMLDMSSDVMIMVMSGSGIGGSLGSM 122
 Oy 120 ILGSTATLAEKACCPVAVMRSKVDQPASDINW---IYVRMTDADPNEAVLEAYAREKTL 176
 Db 123 VMGSVAIVVSHADCPVVVVS--DNHYTEIKTKGPPVVVGSDVDSORATEFAFEAAQA 180

Oy 177 R-----CAPITALGRPEELREIPDGE-----FERRVODWHHRPDVAVYP 217
 Db 181 RGAKLVAHTWMDWQVQASLAGLAAQGEW-EIIEKEQTLLKDRLOPLERFPDVEVM 239
 Oy 218 ITHHTIGARPLADHDERVOALAVIGG-GEAGQALNVGPSGHPVRHAECSLTVR 271
 Db 240 VITDRPVRVRALEDCAHNAQLLVGSHRGGRGMLGSTRALLQSAPCPMVVR 294
 RESULT 14
 AAU62080
 ID AAU62080 standard; protein; 316 AA.
 XX
 AC AAU62080;
 XX
 DT 27-FEB-2002 (first entry)
 XX
 DE Propionibacterium acnes immunogenic protein #22976.
 XX
 KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant.
 XX
 OS Propionibacterium acnes.
 XX
 PN MO200181581-A2.
 XX
 PD 01-NOV-2001.
 XX
 PF 20-APR-2001; 2001WO-US012865.
 XX
 PR 21-APR-2000; 2000US-0199047P.
 PR 02-JUN-2000; 2000US-0208841P.
 PR 07-JUL-2000; 2000US-0216747P.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Skelky YM, Persing DH, Mitcham JL, Wang SS, Bhactia A;
 P1 L'maisonneuve J, Zhang Y, Jen S, Carter D;
 DR N-PSDB; AAS59623.
 XX
 PT Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris.
 XX
 PS Example 1; SEQ ID NO 23275; 1069pp; English.
 XX
 CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
 CC polypeptides. The proteins and their associated DNA sequences are used in
 CC the treatment, prevention and diagnosis of medical conditions caused by
 CC *P. acnes*. The disorders include SAPHO syndrome (synovitis, acne,
 CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
 CC *P. acnes* is also involved in infections of bone, joints and the central
 CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of *P. acnes* in a patient comprises contacting a
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC specific for *P. acnes* proteins. These antibodies can be used to
 CC downregulate expression and activity of *P. acnes* polypeptides and
 CC therefore treat *P. acnes* infections. The antibodies may also be used as
 CC diagnostic agents for determining *P. acnes* presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
 CC this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 CC
 XX Sequence 316 AA;
 SO

	Query Match	10.8%	Score 151;	DB 4;	Length 316;	
	Best Local Similarity	21.8%;	Pred. No. 1.1e-06;			
	Matches	65;	Conservative	46;	Mismatches 135;	Indels 52; Gaps 7
Qy	2	SGRGEPYTKTIIVGIDSGHAAITTAATACGDEAISRVAPLRLVSVIKPTHPSPDD-----	55			
Dd	39	SSKAQRV--IIIVGDSESDGLRAARAYAGSAIKKDADILVHA-----DDAAVAGA	89			
Qy	56	----YDRDLAHASRLREAQSAVEAKLV-----KIETDIPRGPAGEV	95			
Dd	90	MGVVYD-----PTALDAQGVVDVAIHVATERGMDPDRISGEVLGNPAI	136			
Qy	96	LVEASRDAAEMTCVGSVIGRYASSILGSTATLAEAKCPAVM-RSKYDOQASDIINLT	154			
Dd	137	LADSRSADAQLVTLGRRATSGLERMVGGSTSVAVAGSAPRVVVISPSASPDPETGGKKCYA	196			
Qy	155	VRMTPDAFNEAVLEFAAREAKRKORPIIALAGRPBELRIIPDGFEPR---VQMHNHRP	211			
Dd	197	VAVGPQSVGTAVGVGFADRGCKKLAVTVPGNDSEBVNDEALKRLNEVVKPLANKRP	256			
Qy	212	DVRVVPITTHGIARFLADHDREVOLAVGGEAGOLATVSPGHPVFHYNAECSTYL	269			
Dd	257	HVEVETRVLGSEPVDVALDSGVLDVLTGMKKHPILGTAGSGVSRALMAHQSPLAT	314			

RESULT 15	
ABM58599	
ID	ABM58599 standard; protein; 316 AA.
XX	
AC	ABM58599;
XX	
DT	20-OCT-2003 (first entry)
XX	
DE	Propionibacterium acnes predicted ORF-encoded polypeptide #2375.
XX	
KW	Acne vulgaris; antiseborrheic; dermatological; antibacterial;
KW	Immunostimulant; immune response; vaccine.
XX	
OS	Propionibacterium acnes.
XX	
PN	WO200303515-A1.
XX	
XX	24-APR-2003.
XX	
PF	11-OCT-2002; 2002WO-US032727.
XX	
PR	15-OCT-2001; 2001US-00978825.
XX	
PA	(CORI-) CORIXA CORP.
PI	Mitcham JI, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JI,
PI	Zhang Y, Wang S, Jen S, Lodes MW, Benson DR, Jones R, Carter D;
PI	Barth B, Valliave-Douglas J;
XX	
DR	WPI; 2003-381789/36.
XX	
N-PSDB;	ACF64552.
XX	
PT	New Propionibacterium acnes polypeptides and polynucleotides encoding the
PT	polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
PT	or for stimulating an immune response specific for a P. acnes protein.
XX	
PS	Example 1; SEQ ID NO 2375; 1461pp; English.
XX	
CC	The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
CC	encoding a Propionibacterium acnes protein. The invention also relates to
CC	polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
CC	immunogenic fragments of P. acnes polypeptides. The invention
CC	additionally encompasses expression vectors and host cells comprising a
CC	polynucleotide of the invention; antibodies against polypeptides of the
CC	invention; fusion proteins comprising a polypeptide of the invention; a
CC	method for stimulating an immune response specific for a P. acnes
CC	polypeptide and an isolated T cell population comprising T cells prepared

via this method; a vaccine composition (comprising P. acnes polynucleotides, polynucleotides, antibodies, fusion proteins, T cell populations, or antigen-presenting cells that express the polypeptide); a method and kit for detecting or determining the presence or absence of P. acnes in a patient; and a method for inhibiting the development of P. acnes in a patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion proteins, T cell populations or antigen-presenting cells that express the polypeptides are useful for diagnosing, preventing or treating acne vulgaris, or for stimulating an immune response specific for a P. acnes protein. The polynucleotides can also be used as probes or primers for nucleic acid hybridisation. The vaccine composition is useful for the stimulation of an immune response against P. acnes, or for treating acne, and the kit is useful for performing a diagnostic assay. The present sequence represents a polypeptide predicted to be encoded by an ORF (open reading frame) contained within the P. acnes polynucleotides of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp.wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences)

	Query Match	Similarity	10.8%;	Score 151;	DB 6;	length 316;	
	Best Local	Similarity	21.8%;	Pred. 1.1e-06;			
	Matches	Conservative	46;	Mismatches 135;	Indels 52;	Gaps 7	
Qy	2	SGRGEPTKTIIVIGIDGSHAAITAAALMGVDDEAISRAVPLRLVTSVIKPTHPSPDD-----	55				
Db	39	SSKQARTV--IIIVVDGSEDEGLRAARYAMGSAIKRDAOLIVLHV-----DDAAVAGA	89				
Qy	56	-----YDRDLHAERSLREAOFAVEAGLV-----KLETDPGRGAPV	95				
Db	90	MGVYVD-----PTLQDAGGVVDDAIHVATERGMDPRISGEVLVGNPAAI	136				
Qy	96	LVEASRDLEMI CVGSGVIGRYAASSITLGGSTATLEAKHCPVAVM-RSKYDOPASDINMIV	154				
Db	137	LADRSADAQIVLGRATSGLERMFVSGTSAVAVGMSAPRVVISRASTPDTGGKCAVA	196				
Qy	155	VRMTDADPNEVNVLEAAREAKLRQAPILATGCRPELREIPDEGEERR---VQDMHRRP	211				
Db	197	VAVGRQSGVGLTAAGRGPAEDHRRGKLLAVVPPGNDSEVDEALKLRLNVVKKPLANKRP	256				
Qy	212	DVRVYPTITHTGIAFLADHDERVOLAVTGGGEAGQALRLVGPSPHPVRHAECGVLV	269				
Db	257	HVEVETRLVSGEPYDALVDLGGVNDLVTGKKKPHITLQTAGVSRALMAHAGSPAI	314				

Search completed: March 23, 2006, 05:10:51
Job time : 73.3158 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: March 23, 2006, 05:11:18 ; Search time 10.8216 Seconds
(without alignments)
2418.401 Million cell updates/sec

Title: US-10-617-038-27
1397
Sequence: 1 MSGSGEPTMKTIIVGIDGSH.....GPSGHPVFRHAECSVLVRR 272

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1397	100.0	272	2 G70572	hypothetical prote
2	416	29.8	279	2 G70942	hypothetical prote
3	357.5	25.6	268	2 G70645	hypothetical prote
4	286.5	21.4	294	2 A70942	hypothetical prote
5	286	20.5	295	2 C70759	hypothetical prote
6	267	19.1	297	2 F70572	hypothetical prote
7	258.5	18.5	317	2 B70758	hypothetical prote
8	243.5	17.4	288	2 T37029	hypothetical prote
9	230	16.5	283	2 T36962	hypothetical prote
10	224.5	16.1	301	2 T37031	hypothetical prote
11	220	15.7	294	2 T36949	conserved hypotet
12	186	13.3	328	2 T36963	hypothetical prote
13	160.5	11.5	251	2 T37036	hypothetical prote
14	149	10.7	152	2 T36954	hypothetical prote
15	122.5	8.8	174	2 C86231	hypothetical prote
16	119	8.5	170	2 C64366	hypothetical prote
17	116.5	8.3	162	2 F64918	hypothetical prote
18	115.5	8.3	150	2 G84340	hypothetical prote
19	114	8.2	150	2 A75310	hypothetical prote
20	112.5	8.1	139	2 A80454	conserved hypotet
21	112	8.0	270	2 G69469	conserved hypotet
22	109.5	7.8	166	2 B89555	conserved hypotet
23	109	7.8	941	2 A70722	probable gcwb prot
24	107.5	7.7	1997	2 T30874	virginiamycin S sy
25	107	7.7	143	2 AC1497	conserved hypotet
26	106.5	7.6	143	2 B80349	conserved hypotet
27	105.5	7.6	462	2 T28699	hypothetical prote
28	104.5	7.5	147	2 H87082	conserved hypotet
29	103	7.4	143	2 AD1139	conserved hypotet

30	102.5	7.3	271	2 T35065	tryptophan synthas
31	102	7.3	591	2 F64334	acetolactate synth
32	101	7.2	141	2 AH2004	hypothetical prote
33	101	7.2	171	2 H98097	conserved hypotet
34	99.5	7.1	146	2 B70560	hypothetical prote
35	99.5	7.1	155	2 C86634	conserved hypotet
36	99.5	7.1	422	2 G70707	probable puri prot
37	99	7.1	154	2 AD1272	hypothetical prote
38	98.5	7.1	184	2 C69829	conserved hypotet
39	98.5	7.1	406	2 C70640	hypothetical prote
40	98.5	7.1	421	2 E83459	conserved hypotet
41	98	7.0	145	2 A86854	conserved hypotet
42	97	6.9	150	2 F95233	universal stress p
43	97	6.9	152	2 AF1634	hypothetical prote
44	96.5	6.9	190	2 T30796	double-stranded RN
45	96.5	6.9	190	2 G42508	E3L protein - vacc

ALIGNMENTS

```
RESULT 1
G70572      1 100.0%; Score 1397; DB 2; Length 272;
hypothetical protein Rv2624c - Mycobacterium tuberculosis (strain H37Rv)
C:Species: Mycobacterium tuberculosis
C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C:Accession: G70572
R:Coile, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.;
Rajandream, M.A.; Rogers, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.;
Nature 393, 537-544, 1998
A:Authors: Sgares, R.; Sultson, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A>Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: G70572
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-272 <COL>
A:Cross-references: UNIPROT:O06188; UNIPARC:UPI00000D5EAL; GB:Z95387; GB:AL123456; NID:G
A:Experimental source: strain H37Rv
C:Genetic8:
A:Gene: Rv2624c

Query Match      100.0%; Score 1397; DB 2; Length 272;
Best Local Similarity 100.0%; Pred. No. 3.8e-99;
Matches 272; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1  MSGGEPMTKTIIVGIDGSHAITAALMGVDEAISRAVPLRLSVIRTPHSPDDYRDL 60
DB      1  MSGGEPMTKTIIVGIDGSHAITAALMGVDEAISRAVPLRLSVIRTPHSPDDYRDL 60

OY      61  AHAARSIREAOSANEAAGKIVKTEIDI PRGPAGVLYEASRDAMICVSGIGRYASSI 120
DB      61  AHAARSIREAOSANEAAGKIVKTEIDI PRGPAGVLYEASRDAMICVSGIGRYASSI 120

OY      121  LGSTATLAEAKACPVAVMSKVDQASDINMIVVMTDAPDNEAVLEVAAREAKLRQAP 180
DB      121  LGSTATLAEAKACPVAVMSKVDQASDINMIVVMTDAPDNEAVLEVAAREAKLRQAP 180

OY      181  ILALGSRPEELREIPDGEFERRVDWHRRHPDVAVVYITTHTGIAFLADHDERVQLAVI 240
DB      181  ILALGSRPEELREIPDGEFERRVDWHRRHPDVAVVYITTHTGIAFLADHDERVQLAVI 240

OY      241  GGGAGQLARLVGSGHPVFRHAECSVLVRR 272
DB      241  GGGAGQLARLVGSGHPVFRHAECSVLVRR 272

RESULT 2
hypothetical protein Rv2028c - Mycobacterium tuberculosis (strain H37Rv)
C:Species: Mycobacterium tuberculosis
C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
```

C/Accession: C70942
R/Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Felwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A/Authors: Squares, R.; Suleston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A/Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A/Reference number: A70500; MUID:98295987; PMID:9634230
C/Accession: C70942
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-279 <COL>
A/Cross-references: UNIPROT:O53474; UNIPARC:UPI00000D5EDC; GB:AL021899; GB:AL123456; NID:
A/Experimental source: strain H37Rv
C/Genetics:
A/Gene: Rv2028c

Query Match 29.8%; Score 416; DB 2; Length 279;
Best Local Similarity 36.8%; Pred. No. 2e-24;
Matches 98; Conservative 46; Mismatches 116; Indels 6; Gaps 2;

Qy 11 TIIIGIDGSHAAITTAALMGVDEAISRVPRLVSVIKPTSPDDYD---RDLAAERSL 67
Db 9 SIIVGIDGSKPAVQALAAVDEASRDIPRLKLVIAIEPDDGYAAHGAARKLAALAAENAV 68
Qy 68 REAQSAAVEAAKLVKIEITDIPRGAPVLAASRDAMIYGVSGVIGRVAASIIIGSTATE 127
Db 69 RYAFVAEADRPVKEVEITQERPVTSILIRASAAALALVCGAIGVHHFRPERVSTAA 128
Qy 128 LAERAKHCVAVMRSKRVDPADINWIVRMTDAPDNEAVLEAAREALRQAPILALGCR 187
Db 129 LALSAQCPVAIVRPHRVPIGRDAAMIYVEADGSSDIGVLAVAEAERLRDSPRVAVYCR 188
Qy 188 PEELREIPD---GEFERVODMHHRPVRYVPIITHTGIAFLADHDERVOLAVIGGE 244
Db 189 QSGVGTGDDVRASIDRLARNQRIYPRVQSAAVHGBELDYLAGLGRSVHMYVLSASD 248
Qy 245 AGQLARLVGPSGHPVFRHAECSTLVV 270
Db 249 QEHVEQLVGAAPGNNAVLAQAGCTLLVV 274

RESULT 3
G70645
hypothetical protein Rv3134c - Mycobacterium tuberculosis (strain H37Rv)
C/Species: Mycobacterium tuberculosis
C/Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C/Accession: G70645
R/Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Felwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A/Authors: Squares, R.; Suleston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A/Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A/Reference number: A70500; MUID:98295987; PMID:9634230
C/Accession: G70645
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-268 <COL>
A/Cross-references: UNIPROT:P95192; UNIPARC:UPI00000D603B; GB:Z83867; GB:AL123456; NID:9
A/Experimental source: strain H37Rv
C/Genetics:
A/Gene: Rv3134c

Query Match 25.6%; Score 357.5; DB 2; Length 268;
Best Local Similarity 36.5%; Pred. No. 5.5e-20;
Matches 100; Conservative 35; Mismatches 110; Indels 29; Gaps 6;

Qy 10 KTIIVGIDGSHAAITTAALMGVDEAISRVPRLVSVIKPTSPDDYDRDLAAERSLRE 69
Db 8 RAVVVGIDGSRALAAHMAVDEAVNRDIPRLVYVIDPSQLSAAGCGGSAARAAALHD 67
Qy 70 AQSAAVEAAKLVKIEITDIPRGAPVLAASRDAMICVSGVIGRVAASIIIGSTATELA 129

Db 68 ASRKEVETGDPVKIETVLGCRPLTTKMQSRSAAMLCVSGVD-HVRGRGSAATLA 126
Qy 130 EKACHPVAVMRSKVDQPA--SDINWIVRMTDAPDNEAVEYAREAKLRQAPILALGCR 187
Db 127 GSAIACPAAVIVHPSPAEARATTSQVSAVAEV-----DNQVLRHAFEEARLALGVPLRAVA-- 180
Qy 188 PEELREIPD-----EFERRVODMHHRPVRYVPIITHTGIAFLADHDERVOL 237
Db 181 -VHAETPPDVEQGRALAHVLSRLAHMTRLVPEYVDRALAGSACRHLAANAKPGQL 239
Qy 238 AVTGGGAGQLARLVGPSGHPVFRHAECSTLVV 271
Db 240 FVADSHSAHBLCGAYOPG-----CAVLTVR 264

RESULT 4
A70942
hypothetical protein Rv2026c - Mycobacterium tuberculosis (strain H37Rv)
C/Species: Mycobacterium tuberculosis
C/Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C/Accession: A70942
R/Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Felwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A/Authors: Squares, R.; Suleston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A/Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A/Reference number: A70500; MUID:98295987; PMID:9634230
C/Accession: A70942
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-294 <COL>
A/Cross-references: UNIPROT:O53472; UNIPARC:UPI00000D6040; GB:AL021899; GB:AL123456; NID:
A/Experimental source: strain H37Rv
C/Genetics:
A/Gene: Rv2026c

Query Match 21.4%; Score 298.5; DB 2; Length 294;
Best Local Similarity 28.8%; Pred. No. 1.9e-15;
Matches 83; Conservative 53; Mismatches 119; Indels 33; Gaps 8;

Qy 12 TIIIGIDGSHAAITTAALMGVDEAISRVPRLVSVIKPT-----HSPDDYDRDLAH 62
Db 10 ILVGVDSQAQSNAAVAAAREAVWRQLPITLIVAPVVWGVGOLYAMTMOXQ--N 67
Qy 63 AERSIAREQSAVE---AAGLVKIEITDIPRGAPVLAASRDAMIYGVSGVIGRYASS 119
Db 68 AQQVTEQAREALTNLSGESKRPQVHTLVSNVPTLIDASQAAMLVVSGQMGALGR 127
Qy 120 ILGSTATLEAKHCPVAVMRS-KVDQPADINWIVRMTDAPDNEAVEYAREAKLRQ 178
Db 128 ILGSIISIALHHRCPVAIIHSGNATPPSDAP-VLVGIDGSPASAEATLALADEASRR 186
Qy 179 A-----PILAGSRPEELREIPDGEFERVODMHHRPVRYVPIITHTGT 224
Db 187 VDLVAALHAWTDLGFPLGMDWRERERKEBAE--VLAERLGMQEOYPDVVRHSLVCDKP 244
Qy 225 ARFLADHDERVOLAVIGG-GEAGQLARLVGPSGHPVFRHAECSTLVV 271
Db 245 ARWLHSEDAQLVVGSHRGFGSCMLLGSVSAVAHVRIVYIVR 292

RESULT 5
C70759
hypothetical protein Rv2005c - Mycobacterium tuberculosis (strain H37Rv)
C/Species: Mycobacterium tuberculosis
C/Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C/Accession: C70759
R/Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Felwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998

A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: C70759
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-295 <COL>
A:Cross-references: UNIPROT:Q10851; UNIPARC:UPI000013B62F; GB:Z74025; GB:AL123456; NID:G
A:Experimental source: strain H37Rv
C:Genetics:
A:Gene: RV2005c

Query Match 20.5%; Score 286; DB 2; Length 295;
Best Local Similarity 29.6%; Pred. No. 1.7e-14;
Matches 85; Conservative 48; Mismatches 126; Indels 28; Gaps 8;

Qy 12 IIVGIDGSHAITAALWGVDEAISRVPRLVSYTKP-----PTHPSPDDY-----DRDL 60
Db 10 VVVGVDGSLSDPAACWGCATDAAMNIPDLTVHVVNADVATWPPMPPEYETGWQEDBGR 69
Qy 61 AAHRSLSRAGSVAEAGKLVKIEDIPRGAPVLYEASPDAMI CVSGVIGRYSASI 120
Db 70 QIVANAVLAKEAVGADKKL-SVNSLVPSTPVPVPMVEISNEAEVNVLGSSGGRGLARGL 128
Qy 121 LGSTATLAEKACPVAVMRSKVPDQASDINW--IVVMTDAPDNEAVLYEAAREAKLRQ 178
Db 129 LGSVSSSLVRRAGCGVAVIYHSD-DAVIDPQHPAPLVGIDSPVSELATAVAFDEASRG 187
Qy 179 APIALGRRPE-ELREIPDGEPE-----RRVODMHHRPDVRVYPTTHGTGAR 226
Db 188 VELAVHAMSVEVEVELPGLDFSAVQDEALSLAERLKGQBERYDPEVSRVVCDBRPAR 247
Qy 227 FLADHDERVOLAVIGG-GEAGOLARLVGSGHPVFRHAECSVLVVR 272
Db 248 KLVQKASASQLVVVGSHRGGLTGMLDSVSNVLAHARPVIVARQ 294

RESULT 6

F70572
hypothetical protein RV2623 - Mycobacterium tuberculosis (strain H37Rv)

C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C:Accession: F70572
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.;
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: F70572
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-297 <COL>
A:Cross-references: UNIPROT:O06189; UNIPARC:UPI000003326C; GB:Z95387; GB:AL123456; NID:G
A:Experimental source: strain H37Rv
C:Genetics:
A:Gene: RV2623

Query Match 19.1%; Score 267; DB 2; Length 297;
Best Local Similarity 29.3%; Pred. No. 4.9e-13;
Matches 84; Conservative 42; Mismatches 131; Indels 30; Gaps 8;

Qy 12 IIVGIDGSHAITAALWGVDEAISRVPRLVSYTKP-----THPS-----DYDR 58
Db 10 IIVGIDDPAAQVAVRMARDAELKIKLPLTVHAASPEVATWLEVPGLRMQODHGR 69
Qy 59 DLAAHRSLSRAGSVAEAGKLVKIEDIPRGAPVLYEASPDAMI CVSGVIGRYSASI 118
Db 70 HL--IDDAKLVVEQSLRAGP-PTVHSELVPAAPFTLVDSKDVLAIVVVGGLSGRRPG 126
Qy 119 SILGSTATLAEKACPVAVMRSKVD-QPASDINWIVVMTDAPDNEAVLYEAAREAKLR 177

Db 127 RLIGSVSSGLLRHACPVAVIITHEDSVMPHPQAPVLVGDGSSASSELATIAFDEASRR 186
Qy 178 QAPIALGRRPE-ELREIP-----DGEFERRVQDMHHRPDVRVYPTTHGTGA 225
Db 187 NVDLVALHMSDVVSEMPGIDWPATQSMAEQVLAERLAGQERYPVAVITRVVRRQPA 246
Qy 226 RLADHDERVOLAVIGG-GEAGOLARLVGSGHPVFRHAECSVLVVR 271
Db 247 RQLVGRSEBAQLVVVGSRGGRGYAGMLVGSVGETVAQIARPVIVAR 293

RESULT 7

B70758
hypothetical protein RV1996 - Mycobacterium tuberculosis (strain H37Rv)

C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C:Accession: B70758
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.;
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: B70758
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-317 <COL>
A:Cross-references: UNIPROT:Q10862; UNIPARC:UPI000013B4FB; GB:Z74025; GB:AL123456; NID:G
A:Experimental source: strain H37Rv
C:Genetics:
A:Gene: RV1996

Query Match 18.5%; Score 258.5; DB 2; Length 317;
Best Local Similarity 28.3%; Pred. No. 2.4e-12;
Matches 85; Conservative 47; Mismatches 127; Indels 41; Gaps 9;

Qy 12 IIVGIDGSHAITAALWGVDEAISRVPRLVSYTKPSPD-----DYDR----- 59
Db 10 IIVGVDSPGCHTAVEMARDAQOMRVNVALRVQVVPVITTAPEGMAREYSFGQAQREI 69
Qy 60 -----LAHRSLSRAGS-AVEA-AKLVKIEDIPRGAPVLYEASPDAMI CVGS 110
Db 70 VEHSYLVAQAHQIYEQAHKVALLEASSGRAAQITGEVLAHQIVPTLANISQVAMVLYG 129
Qy 111 VGTIRYASSILGSTATLAEKACPVAVMRSKVPDQASDINW-IVVMTDAPDNEAVLEY 169
Db 130 RGQAVAVGALLGSVSSSLVRRHAGPVAVPEE-PRPARPPHAPVVGIDSPFSGLAETI 188
Qy 170 AARAKLRQAPVILA-----LGGPRELEIPDGE---PERRVQDMHHRPD 212
Db 189 AFDEASRRGVDVLVAHMSDGPFLDFPRLLMAPLEMNLEDBQKMLARKLSGQDHYPD 248
Qy 213 VRVYPTTHGTGARFLADHDERVOLAVIGG-GEAGOLARLVGSGHPVFRHAECSVLVVR 271
Db 249 VVHVAKVVCCRPAPRLLELAQTLQVLVVGSHRGGRFGGMHLSVSRVAVNSGAPVIVAR 308

RESULT 8

T37029
hypothetical protein SCU12.10c - Streptomyces coelicolor

C:Species: Streptomyces coelicolor
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: T37029
R:Murphy, L.; Harris, D.; Thomson, N.R.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, August 1999
A:Reference number: Z21619
A:Accession: T37029
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-288 <MOR>
A:Cross-references: UNIPROT:Q9R148; UNIPARC:UPI0000003B34; EMBL:AL109989; PIDN:CAB53422...
A:Experimental source: strain A3(2)

C:Genetics:
A:Gene: SCOEDB:SCJ12.10C

Query Match 17.4%; Score 243.5; DB 2; Length 288;
Best Local Similarity 26.8%; Pred. No. 2.9e-11;
Matches 85; Conservative 36; Mismatches 123; Indels 51; Gaps 9;

QY 12 IIVGIDGSHAAITAAALMGVDEAISRVAFLRVSVI-----KPT-HPSPDD 55
DB 5 MVVGVDSGESSLGAVNDMADEALHEVPLRIYHVRMDRYEGASLARELGKPSGHVTTDD 64
QY 56 YDRDLAAHRSRLREASQAVEAAGLVKLETDIPRGPRAPVLYEASRDAMTCVSGVIGR 115
DB 65 I---LAATRRARRRHHPDLAVTTATTAEPY-----VLLREARNASAVILGRKGE 114
QY 116 YASSILGSTATELEAKACPAVNMRSKVDOPASD---INWIVNMTDPADEAVLEYAR 172
DB 115 LAGLLGSVSLTVATMSDCEVVTTRGDHDDRAAGGRGRIVGVADAP--TAAVRPCE 172
QY 173 EAKLRQAPILA-----LGGRPBELREIPDG-EFERRVODMHHRRPDVAV 215
DB 173 EARRGALADVAMRCPTHDTVHPPLAGTPERLHERAKLEALAD---APADVRL 229
QY 216 YPTTHGTGIAFLADHDERVOLAVYGGGAGQLARLVGSPGHVFRNAECSVLYV 270
DB 230 RRRRTAEGSGRVLSSASHADLVVGRRRPGCFHRLGRVAHTLLHSACEVAVV 284

RESULT 9

T36962
hypothetical protein SCJ1.29c - Streptomyces coelicolor

C/Species: Streptomyces coelicolor
C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C/Accession: T36962
R:Seeger, K.J.; Harris, D.; Thomson, N.R.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, September 1999
A/Reference number: Z21607
A/Accession: T36962
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-263 <SEB>
A/Cross-references: UNIPROT:Q9RIY5; UNIPARC:UP100000DB39B; EMBL:AL109962; PIDN:CAB53147.
A/Experimental source: strain A3(2)
C/Genetics:
A:Gene: SCOEDB:SCJ1.29c

Query Match 16.5%; Score 230; DB 2; Length 283;
Best Local Similarity 30.0%; Pred. No. 3e-10;
Matches 87; Conservative 35; Mismatches 122; Indels 46; Gaps 9;

QY 12 IIVGIDGSHAAITAAALMGVDEAISRVAFLRV-----SVIKPTHPSPDDYDRD 59
DB 5 LVVGVDSGESSLRAVDMADEALHAVPLVWVFGDLMRYEGALALARRGKPSIDMQADD 64
QY 60 LAHAERSLREASQAVEAAG--LVKLETDIPRGPRAPVLYEASRDAMTCVSGVIGRY 116
DB 65 IL-----AAAIRAGRHRHPLV-VTTEVPDEABEHALLCAGRNASMIVGSRGSGI 115
QY 117 ASSILGSTATELEAKHCPAVNMRSKVDOPR--SDINWIVNMTDPADEAVLEYAARE 174
DB 116 ADRLIGSVSRVVAAGSDCPVVVLRGNHNDRAIGGRNRIVVGVEV--SASVRLAEFME 173
QY 175 KLROAPILALGGRPEBELRE-----IPDGFERRVODMHHRR-----PDVRYVPTT 220
DB 174 RLROVPAVAAMCRPAHETIDHPLPAGSPARRRYERAAKLEALAEADPAVDVAVRYTV 233
QY 221 HTGIARFLADHDERVOLAVYGGGAGQLARLVGSPGHVFRNAECSVLYV 270
DB 234 EGPARAVLPAASAEAGLVIGRRGRLGRV---AAVLRHSACPVVVV 279

RESULT 10

T37031

hypothetical protein SCJ12.12c - Streptomyces coelicolor

C/Species: Streptomyces coelicolor
C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C/Accession: T37031
R:Murphy, L.; Harris, D.; Thomson, N.R.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, August 1999
A/Reference number: Z21619
A/Accession: T37031
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-301 <MCR>
A/Cross-references: UNIPROT:Q9RI46; UNIPARC:UP100000DB3BE6; EMBL:AL109989; PIDN:CAB53424.1
A/Experimental source: strain A3(2)
C/Genetics:
A:Gene: SCOEDB:SCJ12.12c

Query Match 16.1%; Score 224.5; DB 2; Length 301;
Best Local Similarity 25.4%; Pred. No. 8.6e-10;
Matches 78; Conservative 42; Mismatches 130; Indels 57; Gaps 8;

QY 10 KTIIVGIDGSHAAITAAALMGVDEAISRVAFLRVSVIKPTHPSPDDYDRDLAA----- 63
DB 3 RTIVGIDGSPESRAAEAMARREGTLRVVRLHWQPV-PP-----MAQAPLLGAE 55
QY 64 -----ERSLREASQAVEAAGLVKLETDIPRGPRAPVLYEASRDAMTCVSGVIGRY 117
DB 56 THQHTERIPDPTAEGRLRHPGVEVTEQATGNPADALAGTLDELVLGSRALSGLT 115
QY 118 SSILGSTATELEAKHCPAVNMR-----SKVDOPASDINWIVNMTDPADE 164
DB 116 GFLVGSYGOSVIATETPTVLVRAQADENHLKDTGISAAITGRFVVVGLDTSPE 175
QY 165 AVLEYAAREAKLQAPI-----LALGSR-EELREIPDGFERRVODMHN 208
DB 176 AVLSFAEEARRRRAPLTAARAMLPESYTYSLAAGDPREBELARAQALGALLPWRE 235
QY 209 RHPDVRYVPTTHGTGIAFLADHDERVOLAVYGGGAGQLARL-----VPSGHPVRHA 263
DB 236 KYPDVEETETCRGSPAEHLIDARDASTLVVG---RRIRSPFGVIGAVAAVAMNA 291
QY 264 EGSVLYV 270
DB 292 TTPVAVV 298

RESULT 11

T36949
conserved hypothetical protein SCJ1.16c - Streptomyces coelicolor

C/Species: Streptomyces coelicolor
C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C/Accession: T36949
R:Seeger, K.J.; Harris, D.; Thomson, N.R.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, September 1999
A/Reference number: Z21607
A/Accession: T36949
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-294 <SEB>
A/Cross-references: UNIPROT:Q9RI28; UNIPARC:UP100000DB38E; EMBL:AL109962; PIDN:CAB53134.1
A/Experimental source: strain A3(2)
C/Genetics:
A:Gene: SCOEDB:SCJ1.16c

Query Match 15.7%; Score 220; DB 2; Length 294;
Best Local Similarity 24.9%; Pred. No. 1.8e-09;
Matches 73; Conservative 43; Mismatches 133; Indels 44; Gaps 6;

QY 12 IIVGIDGSHAAITAAALMGVDEAISRVAFLRV-----SVIKPTHPSPDDYDRD 59
DB 8 LVVGVDSGDSLALDAVDEAQRGLRLVYASLWERYEGALPAMGRSPS-----E 61
QY 60 LAHAERSLREASQAVEAAGLVKLETDIPRGPRAPVLYEASRDAMTCVSGVIGRYASS 119

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 23, 2006, 04:56:43 ; Search time 69.5674 Seconds
(without alignments)
2758.529 Million cell updates/sec

Title: US-10-617-038-27

Perfect score: 1397
Sequence: 1 MSGGGEPTMTITVIGIDSH.....GPSGHPVRHAECSVLVVR 272

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot_05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1397	100.0	272	Q7TY74_MYCBO	Q7EY74 mycobacteri
2	1397	100.0	272	O06188_MYCTU	O06188 mycobacteri
3	416	29.8	279	Q7TZ44_MYCBO	Q7E444 mycobacteri
4	416	29.8	279	O51474_MYCTU	O51474 mycobacteri
5	357.5	25.6	268	Q7TX68_MYCBO	Q7E688 mycobacteri
6	357.5	25.6	268	P95192_MYCTU	P95192 mycobacteri
7	342.5	24.5	252	Q73V37_MYCBO	Q7E337 mycobacteri
8	298.5	21.4	294	Q7Z445_MYCBO	Q7E445 mycobacteri
9	298.5	21.4	294	O53472_MYCTU	O53472 mycobacteri
10	286	20.5	295	Y2005_MYCTU	Y2005 mycobacteri
11	286	20.5	295	Y2028_MYCBO	P64922 mycobacteri
12	277.5	19.9	318	O8FLR9_CORCF	O8FLR9 mycobacteri
13	276.5	19.8	293	Q73261_MYCPA	Q73261 mycobacteri
14	271.5	19.4	294	Q73262_MYCPA	Q73262 mycobacteri
15	267	19.1	297	Q7TY75_MYCBO	Q7EY75 mycobacteri
16	267	19.1	297	O06189_MYCTU	O06189 mycobacteri
17	266	19.0	301	O8NLJ9_CORGL	O8NLJ9 mycobacteri
18	266	19.0	301	O8NLJ9_CORGL	O8NLJ9 mycobacteri
19	260	18.6	294	Q5YV44_MYCBO	Q5YV44 mycobacteri
20	258.5	18.5	317	Y1996_MYCTU	P05457 mycobacteri
21	258.5	18.5	317	Y2019_MYCBO	P05458 mycobacteri
22	254	18.2	296	O5YVE7_NOCFA	O5YVE7 mycobacteri
23	253.5	18.1	296	O5YVE7_NOCFA	O5YVE7 mycobacteri
24	249	17.8	300	O6NEU7_CORDI	O6NEU7 mycobacteri
25	243.5	17.4	280	O9RI48_STRCO	O9RI48 streptomyc
26	242.5	17.4	290	O82NM7_STRAM	O82NM7 streptomyc
27	234	16.8	293	O5YVP7_NOCFA	O5YVP7 mycobacteri
28	231	16.5	303	O4JY55_CORJX	O4JY55 mycobacteri
29	230	16.5	283	O9RIY5_STRCO	O9RIY5 streptomyc
30	227.5	16.3	300	O9K4H8_STRCO	O9K4H8 streptomyc
31	224.5	16.1	301	O9RI46_STRCO	O9RI46 streptomyc

32	224	16.0	281	2	Q7WZ49_9ACTO	Q7WZ49 nonomuraea
33	220	15.7	294	2	Q9RI28_STRCO	Q9RI28 streptomyc
34	210	15.0	295	2	O82NK3_STRAM	O82NK3 streptomyc
35	210	15.0	345	2	O8G3U1_BIFLO	O8G3U1 bifidobacte
36	205	14.7	312	2	O9K4L5_STRCO	O9K4L5 streptomyc
37	202.5	14.5	345	2	Q7JUM6_RHOBA	Q7JUM6 rhodospirill
38	202	14.5	343	2	Q4NL54_9MICC	Q4NL54 arthrobacte
39	191	13.7	144	2	Q4NDF1_9MICC	Q4NDF1 arthrobacte
40	186	13.3	328	2	O9RIY4_STRCO	O9RIY4 streptomyc
41	180.5	12.9	315	2	Q7UER2_RHOBA	Q7UER2 rhodospirill
42	169.5	12.1	289	2	Q4NE71_9MICC	Q4NE71 arthrobacte
43	168.5	12.1	151	2	Q4NGW3_9MICC	Q4NGW3 arthrobacte
44	162.5	11.6	155	2	Q4NL67_9MICC	Q4NL67 arthrobacte
45	160.5	11.5	201	2	O9RI41_STRCO	O9RI41 streptomyc

ALIGNMENTS

RESULT 1

ID	Q7TY74_MYCBO	PRELIMINARY;	PRT;	272 AA.
AC	O7TY74-			
DT	01-OCT-2003 (TREMBLrel. 25, Created)			
DT	01-OCT-2003 (TREMBLrel. 25, Last sequence update)			
DT	01-MAR-2004 (TREMBLrel. 26, Last annotation update)			
DE	Hypothetical protein Mb2657c.			
GN	OrderedlocusNames=MB2657c;			
OS	Mycobacterium bovis.			
OC	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;			
OC	Corynebacterineae; Mycobacteriaceae; Mycobacterium;			
OX	Mycobacterium tuberculosis complex.			
OX	NCBI_TaxID=1765;			
RN	[1]			
RP	NUCLEOTIDE SEQUENCE.			
RC	STRAIN=AF2122/97;			
RC	MEDLINE=22709107; PubMed=12788972; DOI=10.1073/pnas.1130426100;			
RA	Garnier T., Bigmiller K., Camus J.-C., Medina N., Mansoor H.,			
RA	Pryor M., Dutfoy S., Grondin S., Lacroix C., Monsemp C., Simon S.,			
RA	Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,			
RA	Parthill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.,			
RT	"The complete genome sequence of Mycobacterium bovis.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).			
RL	EMBL; BX248343; CAD94842.1; -; Genomic DNA.			
DR	GO; GO:0006950; P:response to stress; IBA.			
DR	InterPro; IPR006015; Usp.			
DR	InterPro; IPR006016; UspA.			
DR	Pfam; PF00582; Usp, 2.			
DR	PRINTS; PR01438; UNVRS1STRESS.			
KW	Complete proteome.			
SQ	SEQUENCE 272 AA; 29400 MW; 3EE3A17B3C5720DB CRC64;			

Query Match	100.0%; Score 1397; DB 2; Length 272;
Best Local Similarity	100.0%; Pred. No. 2.1e-99;
Matches 272; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 MSGGGEPTMTITVIGIDSHAITAALMGVDEAISRAVPLRLTVIKRTHSPDPDYDDL 60
DB	1 MSGGGEPTMTITVIGIDSHAITAALMGVDEAISRAVPLRLTVIKRTHSPDPDYDDL 60
QY	61 AHAERSLNEAQAAYEAAGKLVKIEITDIPRGAPVULVEASRDAMICVSGVIGRYASSI 120
DB	61 AHAERSLNEAQAAYEAAGKLVKIEITDIPRGAPVULVEASRDAMICVSGVIGRYASSI 120
QY	121 LGSRTLEAEAKACPVAVMRKVDQPSADIMWIIVRMTDADNDAVLEYARERKLRAP 180
DB	121 LGSRTLEAEAKACPVAVMRKVDQPSADIMWIIVRMTDADNDAVLEYARERKLRAP 180
QY	181 ILAAGREERERLEIPDGEFERVODMHRHHPDVAVYPTTTGTGARFLADDERVOLAVI 240
DB	181 ILAAGREERERLEIPDGEFERVODMHRHHPDVAVYPTTTGTGARFLADDERVOLAVI 240
QY	241 GGGAAGQLARLVGPSGHPVFRHAECSVLVVR 272

Db 241 GGGAGQIARLVGSPGHVFRHACSVLVVR 272

RESULT 2

ID 006188_MYCTU PRELIMINARY; PRT; 272 AA.

AC 006188; Q7D6V6;
DT 01-JUL-1997 (TREMBLrel. 04, Created)
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
DT 13-SEP-2005 (TREMBLrel. 31, Last annotation update)
DE Hypothetical protein (Universal stress protein family).
GN OrderedLocustNames=MT2699, RV2624c;
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium tuberculosis complex.
OX NCBI_TaxId=1773;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230; DOI=10.1038/31159;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C.M.,
RA Harris D.E., Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III,
RA Tekela F., Badcock K., Basham D., Brown D., Chillingworth T.,
RA Connor R., Davies R.M., Devlin K., Feltwell T., Gentles S., Hamlin N.,
RA Holroyd S., Hornsby T., Jagels K., Krogh J., McLean J., Moule S.,
RA Murphy J.D., Oliver S., Osborne J., Quail M.A., Rajandream M.A.,
RA Rogers J., Rutter S., Seeger K., Skellern B., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrall B.G.,
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."
RL Nature 393:537-544 (1998).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CDC 1551 / Oshkosh;
RX MEDLINE=22206494; PubMed=12218036;
RX DOI=10.1128/JB.184.19.5479-5490.2002;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J.D., DeBoy R.T., Dodson R.J., Gwinn M.L., Haft D.H.,
RA Hickey E.K., Kolonay J.F., Nelson W.C., Umayam L.A., Esmailova M.D.,
RA Salzberg S.L., Delcher A., Uitterlinden T.R., Weidman J.F., Kouri H.M.,
RA Gill J., Mikula A., Bishai W., Jacobs W.R. Jr., Venter J.C.,
RA Fraser C.M.,
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains."
RL J. Bacteriol. 184:5479-5490 (2002).
RX EMBL, BX842580; CAB0618.1; -, Genomic_DNA.
RX EMBL, AE000516; AAK47015.1; -, Genomic_DNA.
DR PIR, G70572; G70572.
DR TIGR, MT2699; -.
DR Tuberculoid; RV2624c; -.
GO GO:0006950; P:response to stress; IEA.
DR InterPro; IPR006015; Usp.
DR InterPro; IPR006016; UspA.
DR Pfam; PF00582; Usp; 2.
DR PRINTS; PR01438; UNVSISTRESS.
KW Complete proteome.
SQ SEQUENCE 272 AA; 29400 MW; 3EE3A17B3C5720D CRC64;

Query Match 100.0%; Score 1397; DB 2; Length 272;
Best Local Similarity 100.0%; Pred. No. 2.1e-99;
Matches 272; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSGGEPTMTKTIIVIGIDGSHAITAAIAGVDEAISRVPRLVSVIKPTHPSPDDYDDL 60
DB 1 MSGGEPTMTKTIIVIGIDGSHAITAAIAGVDEAISRVPRLVSVIKPTHPSPDDYDDL 60
QY 61 AAERSLREAGSAVAAGLVKLTETDIPRGAPVLVEASRDAMIICVSGVIGYASSI 120
DB 61 AAERSLREAGSAVAAGLVKLTETDIPRGAPVLVEASRDAMIICVSGVIGYASSI 120
QY 121 LGSTATLEAKRAHCPVAAMRSKVDPASDINMIIVVMTDAPDNEAVLEVAAREAKLRQAP 180

Db 121 LGSTATLEAKRAHCPVAAMRSKVDPASDINMIIVVMTDAPDNEAVLEVAAREAKLRQAP 180
QY 181 ILALGRRPEELREIPDGEFFERRVODMHRHPDVAVVYPTTHTGTARFLADHDERVOLAVI 240
DB 181 ILALGRRPEELREIPDGEFFERRVODMHRHPDVAVVYPTTHTGTARFLADHDERVOLAVI 240
QY 241 GGGAGQIARLVGSPGHVFRHACSVLVVR 272
DB 241 GGGAGQIARLVGSPGHVFRHACSVLVVR 272

RESULT 3

ID Q7T244_MYCBO PRELIMINARY; PRT; 279 AA.

AC Q7T244;
DT 01-OCT-2003 (TREMBLrel. 25, Created)
DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Hypothetical protein Mb2053c.
GN OrderedLocustNames=Mb2053c;
OS Mycobacterium bovis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium tuberculosis complex.
OX NCBI_TaxId=1765;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AF2123/97;
RX MEDLINE=22709107; PubMed=12788972; DOI=10.1073/pnas.1130426100;
RA Garnier T., Eiglmeier K., Camus J.-C., Medina N., Mansoor H.,
RA Pryor M., Duthey S., Grondin S., Lacroix C., Monempe C., Simon S.,
RA Harris B., Alkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA Parkhill J., Barrall B.G., Cole S.T., Gordon S.V., Hewinson R.G.,
RT "The complete genome sequence of Mycobacterium bovis."
RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882 (2003).
RX EMBL, BX248341; CAB96906.1; -, Genomic_DNA.
DR GO; GO:0006950; P:response to stress; IEA.
DR InterPro; IPR006015; Usp.
DR InterPro; IPR006016; UspA.
DR Pfam; PF00582; Usp; 1.
DR PRINTS; PR01438; UNVSISTRESS.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 279 AA; 29441 MW; 52F240F9FE466AC9 CRC64;

Query Match 29.8%; Score 416; DB 2; Length 279;
Best Local Similarity 36.8%; Pred. No. 7.8e-24;
Matches 98; Conservative 46; Mismatches 116; Indels 6; Gaps 2;

QY 11 TIIVIGIDGSHAITAAIAGVDEAISRVPRLVSVIKPTHPSPDDYD--RDLAHERSL 67
DB 9 SIIVIGIDGSKPAVQAAIAMAVIDEASRDIPRLTLVAIPDDPGYAAHGAARKLAAMNAV 68
QY 68 REAGSAVEAAGKLVKITDIPRGAPVLVEASRDAMIICVSGVIGRYASSILGSTATE 127
DB 69 RYAFVAEADRPKVEEITQERPVTSLIRASAAALVCAGALGVHFRPERVGTAA 128
QY 128 LAERACPCVAAMRSKVDPASDINMIIVVMTDAPDNEAVLEVAAREAKLRQAPILALGCR 187
DB 129 LALAGCPVAIVRRHRPIRGDAIYVEADGSDIGVLGAWAERLADSPRVVYTCR 188
QY 188 PEELREIPD--GEFFERRVODMHRHPDVAVVYPTTHTGTARFLADHDERVOLAVIGGE 244
DB 189 QSGVGTGDVDVRAISLDRLARWQGRYPDVAVVQSAVAGELLDYLAGRGSVHVVLSASD 248
QY 245 AGQALRLVGPBGHVPFRHACSVLV 270
DB 245 QEHVEQLVGAAGNAVLQEAAGCTLLVV 274

RESULT 4

ID 053474_MYCTU PRELIMINARY; PRT; 279 AA.

```

AC 053474; Q7D7L5;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein (Universal stress protein family).
GN OrderedLocustNames=MT2087, Rv2028c;
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium tuberculosis complex.
OX NCBI_TaxID=1773;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=H37Rv;
RX MEDLINE=98295987; PubMed=9634230; DOI=10.1038/31159;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C.M.,
RA Harris D.E., Gordon S.V., Eigmeier K., Gas S., Barry C.E. III,
RA Tekle A.F., Badcock K., Basham D., Brown D., Chillingworth T.,
RA Connor R., Davies R.M., Devlin K., Felwell T., Gentles S., Hamlin N.,
RA Holroyd S., Hornby T., Jagels K., Kesteven A., McLean J., Moule S.,
RA Murphy L.D., Oliver S., Osborne J., Quail M.A., Rajandream M.A.,
RA Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulten J.E., Taylor K., Whitehead S., Barrall B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."
RL Nature 393:537-544 (1998).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CDC 1551 / Oshkosh;
RX MEDLINE=22206494; PubMed=12218036;
RX DOI=10.1186/1471-2191-5479-5490.2002;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J.D., DeBoy R.T., Dodson R.C., Gwinn M.L., Haft D.H.,
RA Hickey B.K., Kolonay J.F., Nelson W.C., Umayam L.A., Brmlaeva M.D.,
RA Salzberg S.L., Delcher A., Utterback T.R., Melman J.F., Khouri H.M.,
RA Gill J., Mikula A., Bishai W., Jacobs W.R. Jr., Venter J.C.,
RA Fraser C.M.;
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains."
RL J. Bacteriol. 184:5479-5490 (2002).
DR EMBL; BX942578; CA17242.1; -; Genomic DNA.
DR EMBL; AE000516; AAK46366.1; -; Genomic DNA.
DR PIR; C70942; C70942.
DR TIGR; MT2087; -.
DR Tuberculist; Rv2028c; -.
DR GO; GO:0006950; P:response to stress; IEA.
DR InterPro; IPR006015; Unp.
DR InterPro; IPR006016; UnpA.
DR Pfam; PF00582; Unp; 1.
DR PRINTS; PR01438; UNYRS1STRESS.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 279 AA; 29441 MW; 52F240F9FE466AC9 CRC64;

```

```

Query Match 29.8%; Score 416; DB 2; Length 279;
Best Local Similarity 36.8%; Pred. No. 7.8e-24;
Matches 98; Conservative 46; Mismatches 116; Indels 6; Gaps 2;
QY 11 TTIVGIDGSHAATTAALMGVDEAISRAVPLRLVSVIKPTHPSPDDYD---RDLNAEBSL 67
DB 9 SIVVGIDGSKPAVQALMAVDEASRDIPRLRLVLAIEPDDPGYAAHGAARKLAANAV 68
QY 68 REAASAVNAAGKLVKIEETDIPRGPAVLVEASRDAMI CVSGVIGRVAASIIIGSTATE 127
DB 69 RYATPAVDAARPVYVEVEIQERPVTSILIRASAAALVCAIGALGVNHFRRPVGSTAA 128
QY 128 LAEYAHCPVAVMRSKVDPADINMIVVMTDAPDNEAVLEAAREAKLRQAPILALGGR 187
DB 129 LALSAQCVAIVRPHRVIGRDAAMI VVEADGSSDIGTLGAVMAEARLRDSPRVAVTGR 188
QY 188 PEELREITD---GEERRVODMHNHPRVRYVPIITHTGIGARPLADHDERVOLAVIGGE 244
DB 189 QSGVGDGDDVRAISLRWLAHQPRYRVDVQVSAAVHGBELLDYLAGLRSVHVVVLSASD 248

```

```

QY 245 AGQIARLVGSPGHPVFRHAECSVLV 270
DB 249 QEHVEQLVAGPAGNAVLOEAGCTLLVV 274
RESULT 5
Q7TX68_MYCBO
ID Q7TX68_MYCBO PRELIMINARY; PRT; 268 AA.
AC Q7TX68;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein MD3158c.
GN OrderedLocustNames=MD3158c;
OS Mycobacterium bovis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium tuberculosis complex.
OX NCBI_TaxID=1765;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AF2122/97;
RX MEDLINE=22709107; PubMed=12788972; DOI=10.1073/pnas.1130426100;
RA Garnier T., Eigmeier K., Camus J.-C., Medina N., Mansoor H.,
RA Pryor M., Duthey S., Grondin S., Lacroix C., Monsemp C., Simon S.,
RA Harris B., Alkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA Parkhill J., Barrall B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RT "The complete genome sequence of Mycobacterium bovis."
RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882 (2003).
DR EMBL; BX248345; CAD95250.1; -; Genomic DNA.
DR GO; GO:0006950; P:response to stress; IEA.
DR InterPro; IPR006015; Unp.
DR InterPro; IPR006016; UnpA.
DR Pfam; PF00582; Unp; 1.
DR PRINTS; PR01438; UNYRS1STRESS.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 268 AA; 28008 MW; E82CB4F9C9D602AE CRC64;

```

```

Query Match 25.6%; Score 357.5; DB 2; Length 268;
Best Local Similarity 36.5%; Pred. No. 2.4e-19;
Matches 100; Conservative 35; Mismatches 110; Indels 29; Gaps 6;
QY 10 KTIIVGIDGSHAATTAALMGVDEAISRAVPLRLVSVIKPTHPSPDDYDRDLNAEBSLRE 69
DB 8 RAVVVGIDGSRRAATTAALMAVDEAVNRDIPRLRLVVIDPQLSAAGCGGQSAARALHD 67
QY 70 AQSAYEAAAGKLVKIEETDIPRGPAVLVEASRDAMI CVSGVIGRVAASIIIGSTATE 129
DB 68 ASRVEAYTGPVKIETEVLCGRPLTKMQSRSAAMLCVSGVGLD-HVGRGRGSVVAATLA 126
QY 130 EKACHCPVAVMRSKVDPDA--SDINMIVVMTDAPDNEAVLEAAREAKLRQAPILALGGR 187
DB 127 GSAICPVAIVHPSAERPTISQVSAVAEV---DNGVLRHARPEEARLRQVPLRAVA-- 180
QY 188 PEELREITD-----EFERRVODMHNHPRVRYVPIITHTGIGARPLADHDERVOL 237
DB 181 -VHAETRPDVEQSRILAHVHLSRLAHMTRLYBEVAVDAIAGSACRHLAANAKGQL 239
QY 238 AVTGGAGQIARLVGSPGHPVFRHAECSVLV 271
DB 240 FVADSHSAHELCAYQPG-----CAVLTVR 264
RESULT 6
P95192_MYCTU
ID P95192_MYCTU PRELIMINARY; PRT; 268 AA.
AC P95192; Q7D624;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein (Universal stress protein family).
GN OrderedLocustNames=MT3220, Rv3134c;
OS Mycobacterium tuberculosis.

```

OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacteriineae; Mycobacteriaceae; Mycobacterium;
 OC Mycobacterium tuberculosis complex.
 NC NCB1_TaxID=1773;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=H37Rv.
 RX MEDLINE=98295987; PubMed=9634230; DOI=10.1038/31159;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C.M.,
 RA Harris D.E., Gordon S.V., Eigmeier K., Gae S., Barry C.E. III,
 RA Tekala F., Badcock K., Basham D., Brown D., Chillingworth T.,
 RA Connor R., Davies R.M., Devlin K., Fellwell T., Gentles S., Hamlin N.,
 RA Holroyd S., Hornby T., Jagels K., Krogh A., McLean J., Moule S.,
 RA Murphy L.D., Oliver S., Osborne J., Quail M.A., Rajandream M.A.,
 RA Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 RA Sulston J.E., Taylor K., Whitehead S., Barrett B.G.;
 RT "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence."
 RL Nature 393:537-544 (1998).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=CDC 1551 / Oshkosh;
 RX MEDLINE=22206494; PubMed=12218036;
 DOI=10.1128/JB.184.19.5479-5490.2002;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J.D., DeBoy R.T., Dodson R.J., Gwinn M.L., Haft D.H.,
 RA Hickey E.K., Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D.,
 RA Salzberg S.L., Delcher A.L., Uterback T.R., Weisman J.F., Khouri H.M.,
 RA Gill J., Mikula A., Blahni W., Jacobs W.R. Jr., Venter J.C.,
 RA Fraser C.M.;
 RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains."
 RL J. Bacteriol. 184:5479-5490 (2002).
 DR EMBL; BX842582; CAB06280.1; -: Genomic DNA.
 DR PIR; G70645; G70645.
 DR TIGR; MT3220; -:
 DR TubercuList; RV3134C; -;
 DR GO; GO:0006950; P:response to stress; IEA.
 DR InterPro; IPR006015; USP.
 DR InterPro; IPR006016; USPA.
 DR Pfam; PF00582; USP; 1.
 DR PRINTS; PR01438; UNYRSLSSTRESS.
 DR Complete proteome; Hypothetical protein.
 KW SEQUENCE 268 AA; 28008 MW; EE2CB4FC9C602AE CRC64;
 SQ
 Query Match 25.6%; Score 357.5; DB 2; Length 268;
 Best Local Similarity 36.5%; Pred. No. 2,4e-19;
 Matches 100; Conservative 35; Mismatches 110; Indels 29; Gaps 6;
 QY 10 KTIIVGIDGSHAAITTAALMGVDEAISRVAVPLRLVSVIKPTHSPDDYDRDLAAHRSURE 69
 DB 8 RAVVVGIDGSRRAATTAALMAVDEAVNRDIPRLVTVVIDPQLSAAGEGCGSAAARALHD 67
 QY 70 AQSVAEAGKLVKIEITDIPRGPAGVLYEASRDAMICVSGVIGRYASSILGSTATELA 129
 DB 68 ASRKEVAAGQPVKLETVLCRPLTKMQESRSAAMLCVSGGLD-HVRGGRGSVAATLA 126
 QY 130 EKACPVAVMSKVDQPA--SDIMNIVVMTDAPDNEAVLEAYAREALRQAPITALAGCR 187
 DB 127 GSALCPVAVIVHPSREPAITTSQVSAVVAE---DNGVVLKHAPEEARLRGVPLRAVA-- 180
 QY 188 PEELREIPDG-----EFERRVODMHHRHDPVRYVPTTHTGIRFLADHDERVOL 237
 DB 181 -VHAETPTDVEQSGRLAHVHLSRLAHMTRLYPEVRVDRALVAGSACRHLAAAKKPGQL 239
 QY 238 AVTGGGAGQTLARLVGSGHVFRRHAECSVLVVR 271
 DB 240 FVADSHSHELCGAYQPG-----CAVLTVR 264

ID Q73V37_MYCPA PRELIMINARY; PRT; 252 AA.
 AC Q73V37;
 DT 05-JUL-2004 (TREMBlrel. 27, Created)
 DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
 DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
 DE Hypothetical protein.
 GN OrderedLocustNames=MAP3179c;
 OS Mycobacterium paratuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacteriineae; Mycobacteriaceae; Mycobacterium;
 OC Mycobacterium avium complex (MAC).
 NC NCB1_TaxID=1770;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=K10;
 RA Li L., Bannantine J., Zhang Q., Amonsin A., Alt D., Kapur V.;
 RL Submitted (SFP-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE017238; AAS05727.1; -: Genomic DNA.
 DR GO; GO:0006950; P:response to stress; IEA.
 DR InterPro; IPR006016; USPA.
 DR Pfam; PF00582; USP; 1.
 DR Complete proteome; Hypothetical protein.
 KW SEQUENCE 252 AA; 26447 MW; FFDAB63EB3EA80E1 CRC64;
 SQ
 Query Match 24.5%; Score 342.5; DB 2; Length 252;
 Best Local Similarity 36.2%; Pred. No. 3,2e-18;
 Matches 98; Conservative 41; Mismatches 89; Indels 43; Gaps 10;
 QY 12 IIVGIDGSHAAITTAALMGVDEAISRVAVPLRLVSVIKPTHSPDDY--DRDLAAHRSURE 69
 DB 10 IVGVDDSRRAAMAAVAVAIDAVGRDIPRLVTVVIDP-HGAPGSGHPTRLAAARALAD 68
 QY 70 AQSVAEAGKLVKIEITDIPRGPAGVLYEASRDAMICVSGVIGRYASSILGSTATELA 129
 DB 69 AHRVDAFAPQPVKVEITILWGNTAFKLEOSRSAAVMLCVQIGL-NHACGCPALATSLV 127
 QY 130 EKACPVAVMSKVDQPA--ASDIMNIVVMTDAPDNEAVLEAYAREALRQAPITALAGCR 188
 DB 128 RSLACPVAVVOQAPSLPAAARVSGVAEV---DNGTVLHHAPEEARLRGVLCVAVNP 183
 QY 189 EELREIPDGSEFERRVODMHHRHDPVRYVPTTHTGIRFL-ADHDERVOLAVIGGGAQ 247
 DB 184 ARV-----ELERRLAMKMLRPVQAESAVLTVGSGVGHRLADH-----RAGR 225
 QY 248 L-----ARLVGSPGHVFRHAECSVLVVR 271
 DB 226 LVTVDAYRABALCHAGH-----SVLAVR 248

RESULT 8
 Q77245_MYCBO
 ID Q77245_MYCBO PRELIMINARY; PRT; 294 AA.
 AC Q77245;
 DT 01-OCT-2003 (TREMBlrel. 25, Created)
 DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
 DE Hypothetical protein Mb2051c.
 GN OrderedLocustNames=Mb2051c;
 OS Mycobacterium bovis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacteriineae; Mycobacteriaceae; Mycobacterium;
 OC Mycobacterium tuberculosis complex.
 NC NCB1_TaxID=1765;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=AF2122/97;
 RX MEDLINE=22709107; PubMed=12788972; DOI=10.1073/pnas.1130426100;
 RA Garnier T., Eigmeier K., Camus J.-C., Medina N., Mansoor H.,
 RA Pryor M., Dutfoy S., Gordin S., Lacroix C., Monempe C., Simon S.,
 RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
 RA Parkhill J., Barrett B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
 RT "The complete genome sequence of Mycobacterium bovis."
 RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882 (2003).

RESULT 7
 Q73V37_MYCPA

DR EMBL, BX248341; CAD96904.1; -; Genomic_DNA.
 DR GO; GO:0006950; P:response to stress; IEA.
 DR InterPro; IPR006015; USP.
 DR InterPro; IPR006016; USPA.
 DR Pfam; PF00582; USP; 2.
 DR PRINTS; PR01438; UNRSLSSTRESS.
 DR Complete proteome; Hypothetical protein.
 KW SEQUENCE 294 AA; 31456 MW; F0627782110F7142 CRC64;
 SQ

Query Match 21.4%; Score 298.5; DB 2; Length 294;
 Best Local Similarity 28.8%; Pred. No. 9.3e-15;
 Matches 83; Conservative 53; Mismatches 119; Indels 33; Gaps 8;

QY 12 IIVGIDGSHAATTAALMGVDEAISRVAFLRLSVIKPT-----HSPDDYDRDLAH 62
 DB 10 ILVGVDSAGQNAVAAMAAAREAVNRQLPTLTHVAAPVVGMPVQGLYANTEMQKD--N 67
 QY 63 AERSLREAGSAVE---AAGKLVKIETDIPRGPAGPVLEASRDAMTCVSGVIGRYASS 119
 DB 68 AQQVIEQAREALTNLSLGSKPQVHTLVFSNVVPTLLIDASQAMLMVVGSGQMGALGRL 127
 QY 120 ILGSTATLAKKACFPVAVNRS-KVDOPASDINMIVRMDAPDNEAVLEVAAREAKLRQ 178
 DB 128 ILGSIISTALLHARCPVALIHSGNGATPDSAP-VLVGIDGSPASEATATAFDEARRR 186
 QY 179 A-----PIALGGRPEELREIPDGEFERRVODMHRHPDVRVYPITHTGI 224
 DB 187 VDLVALHAWTDLGMFPVLGMDWREREREAE--VLAERLACMOQEQYDVRVHRSLVCDKP 244
 QY 225 ARFLADHDERVOLAVTGG-GEAGQLARLVGSPGHVFRHAECSVLVVR 271
 DB 245 ARWLLEHSEQAQLVVGSHGRGFGSGMLGSSVSAVHVRIPVIVVR 292

RESULT 9
 ID 053472 MYCTU PRELIMINARY; PRT; 294 AA.
 AC 053472; Q7D7L7;
 DT 01-JUN-1998 (TReMBLrel. 06, Created)
 DT 01-JUN-1998 (TReMBLrel. 06, Last sequence update)
 DT 13-SEP-2005 (TReMBLrel. 31, Last annotation update)
 DE Hypothetical protein (Universal stress protein family).
 GN OrderedLocustNames=MT2085, RV2026c;
 OS Mycobacterium tuberculosis.
 OC Bacteria; Actinobacteria; Actinomycetales;
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
 OC Mycobacterium tuberculosis complex.
 OX NCBI_TaxID=1773;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=H37RV;
 RX MEDLINE=98295987; PubMed=9634230; DOI=10.1038/31159;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C.M.,
 RA Harris D.E., Gordon S.V., Eigmeier K., Gae S., Barry C.E. III,
 RA Tekala F., Badcock K., Basham D., Brown D., Chillingworth T.,
 RA Connor R., Davies R.M., Devlin K., Felwell T., Gentles S., Hamlin N.,
 RA Holroyd S., Hornsby T., Jagers K., Krogh A., McLean J., Moule S.,
 RA Murphy L.D., Oliver S., Osborne J., Quail M.A., Rajandream M.A.,
 RA Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 RA Sultston J.E., Taylor K., Whitehead S., Barrett B.G.;
 RT "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence."
 RL Nature 393:537-544 (1998).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=CDC 1551 / Oshkosh;
 RX MEDLINE=22206494; PubMed=12218036;
 RX DOI=10.1128/DB.184.19.5479-5490.2002;
 RA Flaischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J.D., DebRoy R.T., Dodson R.J., Gwin M.L., Haft D.H.,
 RA Hickey E.K., Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D.,
 RA Salzberg S.L., Delcher A., Uitterback T.R., Weisman J.F., Kouri H.M.,
 RA Gill J., Mikula A., Biehni W., Jacobs W.R. Jr., Venter J.C.,

RA Fraser C.M.;
 RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains."
 RL J. Bacteriol. 184:5479-5490 (2002).
 DR EMBL, BX842578; CA417240.1; -; Genomic_DNA.
 DR EMBL, AE000516; AKA46364.1; -; Genomic_DNA.
 DR PIR; A70942; A70942.
 DR TIGR; MT2085; -.
 DR Tuberculist; RV2026c; -.
 DR GO; GO:0006950; P:response to stress; IEA.
 DR InterPro; IPR006015; USP.
 DR InterPro; IPR006016; USPA.
 DR Pfam; PF00582; USP; 2.
 DR PRINTS; PR01438; UNRSLSSTRESS.
 DR Complete proteome; Hypothetical protein.
 KW SEQUENCE 294 AA; 31456 MW; F0627782110F7142 CRC64;
 SQ

Query Match 21.4%; Score 298.5; DB 2; Length 294;
 Best Local Similarity 28.8%; Pred. No. 9.3e-15;
 Matches 83; Conservative 53; Mismatches 119; Indels 33; Gaps 8;

QY 12 IIVGIDGSHAATTAALMGVDEAISRVAFLRLSVIKPT-----HSPDDYDRDLAH 62
 DB 10 ILVGVDSAGQNAVAAMAAAREAVNRQLPTLTHVAAPVVGMPVQGLYANTEMQKD--N 67
 QY 63 AERSLREAGSAVE---AAGKLVKIETDIPRGPAGPVLEASRDAMTCVSGVIGRYASS 119
 DB 68 AQQVIEQAREALTNLSLGSKPQVHTLVFSNVVPTLLIDASQAMLMVVGSGQMGALGRL 127
 QY 120 ILGSTATLAKKACFPVAVNRS-KVDOPASDINMIVRMDAPDNEAVLEVAAREAKLRQ 178
 DB 128 ILGSIISTALLHARCPVALIHSGNGATPDSAP-VLVGIDGSPASEATATAFDEARRR 186
 QY 179 A-----PIALGGRPEELREIPDGEFERRVODMHRHPDVRVYPITHTGI 224
 DB 187 VDLVALHAWTDLGMFPVLGMDWREREREAE--VLAERLACMOQEQYDVRVHRSLVCDKP 244
 QY 225 ARFLADHDERVOLAVTGG-GEAGQLARLVGSPGHVFRHAECSVLVVR 271
 DB 245 ARWLLEHSEQAQLVVGSHGRGFGSGMLGSSVSAVHVRIPVIVVR 292

RESULT 10
 ID Y2005 MYCTU STANDARD; PRT; 295 AA.
 AC P64921; Q10851;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 25-OCT-2004 (Rel. 45, Last sequence update)
 DT 10-MAY-2005 (Rel. 47, Last annotation update)
 DE Hypothetical protein RV2005c/MT2061.
 GN OrderedLocustNames=RV2005c, MT2061; ORFNames=MTCY39.12;
 OS Mycobacterium tuberculosis.
 OC Bacteria; Actinobacteria; Actinomycetales;
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
 OC Mycobacterium tuberculosis complex.
 OX NCBI_TaxID=1773;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=H37RV;
 RX MEDLINE=98295987; PubMed=9634230; DOI=10.1038/31159;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C.M.,
 RA Harris D.E., Gordon S.V., Eigmeier K., Gae S., Barry C.E. III,
 RA Tekala F., Badcock K., Basham D., Brown D., Chillingworth T.,
 RA Connor R., Davies R.M., Devlin K., Felwell T., Gentles S., Hamlin N.,
 RA Holroyd S., Hornsby T., Jagers K., Krogh A., McLean J., Moule S.,
 RA Murphy L.D., Oliver S., Osborne J., Quail M.A., Rajandream M.A.,
 RA Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 RA Sultston J.E., Taylor K., Whitehead S., Barrett B.G.;
 RT "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence."
 RL Nature 393:537-544 (1998).
 RN [2]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

Db 243 RPARQLIQSSSAQLTVVGSHGSGGLAGTLGVSNAVHSHVMPVIAR 292

RESULT 15

07TY75_MYCBO PRELIMINARY; PRT; 297 AA.
ID 07TY75_MYCBO PRELIMINARY; PRT; 297 AA.
AC 07TY75;
DT 01-OCT-2003 (TREMBlrel. 25, Created)
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Hypothetical protein TB31.7.
GN Name=TB31.7; OrderedLocusName=Mb2656;
OS Mycobacterium bovis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium tuberculosis complex.
CX NCBI_TaxID=1765;
NX 11
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AF2122/97;
RX MEDLINE=22709107; PubMed=12788972; DOI=10.1073/pnas.1130426100;
RA Garnier T., Siglmeier K., Camus J.-C., Medina N., Mansoor H.,
RA Pryor M., Duchoy S., Grondin S., Lacroix C., Monsempé C., Simon S.,
RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RT "The complete genome sequence of Mycobacterium bovis."
RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
DR EMBL: BX248343; CAD94841.1; -: Genomic DNA.
DR GO: 0006950; P:response to stress; TGA.
DR InterPro: IPR006015; UBP.
DR InterPro: IPR006016; UBP.
DR Pfam: PF00582; UBP; 2.
DR PRINTS: PR01438; UNVRSLSTRESS.
KW Complete proteome.
SQ SEQUENCE 297 AA; 31652 MW; A590F7058D1E8695 CRC64;

Query Match 19.1%; Score 267; DB 2; Length 297;
Best Local Similarity 29.3%; Pred. No. 2.5e-12;
Matches 84; Conservative 42; Mismatches 131; Indels 30; Gaps 8;

QY 12 IIVGDSHAITATLWGVDEISRAVPLRLVSVIKP-----THRSP-----DQYDR 58
DB 10 IIVGIDSPAAQVAVRWAPARDAELRKIPLTIVHAVSPVATWLEVPPLPGVLRWQDHR 69
QY 59 DLAAERSLREASVBAAGKLVKLETDI PRGPAGPVLVEASRDALMI CVGSVGIGRYAS 118
DB 70 HL--IDDLKLVVEQASLRAGP-PTVHSEIVPAAAVPTLVMSKDAVLAVVGCISGRWPG 126
QY 119 SILGSTATLEAKAHCPVAVNRSKYD-QPASDINMIVVMTDAPDNEAVLEVAAREAKLR 177
DB 127 RLGSVSSGLRHAHCPVITIHEDBSVMPHPOQAPVLGVDSGSASSELATATAIAFDEASRR 186
QY 178 QAPITALGGRPE-ELAREIP-----DGEFERVQDWHHHPDVRVYPITHTGIA 225
DB 187 NVDLVALHAWSDVDSEWPGIDWPATQSMAEVLAEIRLAGWQERYPNVAITRVVVRDQPA 246
QY 226 RFLADHDERVOLAVIGG--GEAGQLARLVGPSGHPVFRHAECSTLVVR 271
DB 247 RQLVORSEEAQLVVVGSRGGRGAGMLVGSVGETVTAQLARTPTVIYAR 293

Search completed: March 23, 2006, 05:24:26
Job time : 70.5674 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: March 23, 2006, 05:24:53 ; Search time 18.5513 Seconds
(without alignments)
1212.194 Million cell updates/sec

Title: US-10-617-038-27
Perfect score: 1397
Sequence: 1 MSGRGPMTKTIIVIGIDSH.....GPSGHVPRHAECSVLVRR 272

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 10%
Listing first 45 summaries

Database : Issued Patents_AA:
1: /cgn2_6/ptodata/1/1aa/5_COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/6_COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/H_COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/PCetus_COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/RE_COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	266	19.0	301	2 US-09-605-703B-1870	Sequence 1870, Ap
2	112	8.0	189	2 US-09-134-001C-3795	Sequence 3795, Ap
3	105.5	7.6	613	2 US-09-758-759-151	Sequence 151, App
4	105	7.5	151	2 US-09-107-532A-5606	Sequence 5606, Ap
5	103.5	7.4	430	2 US-09-902-540-12843	Sequence 12843, A
6	100.5	7.2	1028	2 US-09-328-352-5749	Sequence 5749, Ap
7	98.5	7.1	542	2 US-09-252-991A-21975	Sequence 21975, A
8	98.5	7.1	735	2 US-09-252-991A-30569	Sequence 30569, A
9	98	7.0	657	2 US-09-489-039A-10221	Sequence 10221, A
10	97.5	7.0	715	2 US-09-902-540-12939	Sequence 12939, A
11	97	6.9	148	2 US-09-107-433-3293	Sequence 3293, Ap
12	97	6.9	150	2 US-09-583-110-3588	Sequence 3588, Ap
13	97	6.9	161	2 US-09-107-532A-3842	Sequence 3842, Ap
14	97	6.9	420	2 US-09-252-991A-30438	Sequence 30438, A
15	94	6.7	157	2 US-09-328-352-7302	Sequence 7302, Ap
16	94	6.7	800	2 US-09-252-991A-21975	Sequence 21975, A
17	93.5	6.7	531	2 US-09-252-991A-17650	Sequence 17650, A
18	93.5	6.7	695	2 US-09-252-991A-19147	Sequence 19147, A
19	93	6.7	485	2 US-09-489-039A-10554	Sequence 10554, A
20	93	6.7	565	2 US-09-252-991A-23494	Sequence 23494, A
21	92.5	6.6	366	2 US-09-252-991A-32542	Sequence 32542, A
22	92	6.6	367	2 US-09-902-540-14540	Sequence 14540, A
23	92	6.6	844	2 US-09-252-991A-27184	Sequence 27184, A
24	91.5	6.5	331	2 US-09-252-991A-29383	Sequence 29383, A
25	91.5	6.5	452	2 US-09-252-991A-18948	Sequence 18948, A
26	90.5	6.5	431	2 US-09-252-991A-20267	Sequence 20267, A
27	90	6.4	467	2 US-09-252-991A-23045	Sequence 23045, A

28	89.5	6.4	478	2 US-09-252-991A-31866	Sequence 31866, A
29	89.5	6.4	484	2 US-09-252-991A-26896	Sequence 26896, A
30	89.5	6.4	585	2 US-09-252-991A-30045	Sequence 30045, A
31	89.5	6.4	615	2 US-09-949-016-11320	Sequence 11320, A
32	89.5	6.4	684	2 US-09-252-991A-27004	Sequence 27004, A
33	89.5	6.4	1388	2 US-09-252-991A-20237	Sequence 20237, A
34	89.5	6.4	1953	2 US-09-917-254-92	Sequence 92, Appl
35	89	6.4	512	2 US-09-902-540-15789	Sequence 15789, A
36	88.5	6.3	1421	2 US-09-335-409-2	Sequence 2, Appl1
37	88.5	6.3	1421	2 US-09-568-102-2	Sequence 2, Appl1
38	88.5	6.3	1421	2 US-09-568-969-2	Sequence 2, Appl1
39	88.5	6.3	1421	2 US-09-568-480-2	Sequence 2, Appl1
40	88.5	6.3	1421	2 US-09-568-486-2	Sequence 2, Appl1
41	88.5	6.3	1421	2 US-09-568-472-2	Sequence 2, Appl1
42	88.5	6.3	1421	2 US-09-567-899-2	Sequence 2, Appl1
43	88.5	6.3	1421	2 US-10-014-717-2	Sequence 2, Appl1
44	88	6.3	403	2 US-09-489-039A-11881	Sequence 11881, A
45	87.5	6.3	547	2 US-09-252-991A-18576	Sequence 18576, A

ALIGNMENTS

```
RESULT 1
US-09-605-703B-1870
; Sequence 1870, Application US/09605703B
; Patent No. 5962989
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Krieger, Burkhard
; APPLICANT: Schröder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauser, Gregor
; TITLE OF INVENTION: CORNYNEBACTERIUM GLUTAMICUM GENES ENCODING NOVEL
; FILE OF INVENTION: PROTEINS
; FILE REFERENCE: BG1-129CP
; CURRENT APPLICATION NUMBER: US/09/605, 703B
; PRIOR FILING DATE: 2000-06-27
; PRIOR APPLICATION NUMBER: 60/142,764
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: 60/152,318
; PRIOR FILING DATE: 1999-09-03
; NUMBER OF SEQ ID NOS: 2934
; SEQ ID NO 1870
; LENGTH: 301
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-605-703B-1870
Query Match 19.0%; Score 266; DB 2; Length 301;
Best Local Similarity 28.8%; Pred. No. 7.7e-21;
Matches 85; Conservative 45; Mismatches 11; Indels 34; Gaps 8;
QY 8 TMKTIIVIGIDSHAAITPAALMGVDEAISRAVPLRLVSVIKETPHS-----PDDYDR 58
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 3 TEDIIVAAVGVSDASKQAVRMAANTANKRGIPRLAS--SYTWPOFYAEGMVPPOFLFD 60
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 59 DL-AHARSRLREKQSAVEAAGKVKITDTPREGAGVLYEASDAMICVSGICIRYA 117
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 61 DLOEALAEKINEARDIAHEVAPEIKIGHTIAEGSPIDMLLEMSDATTIVGSGLS 120
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 118 SSIIIGSTATLEAAGACGAVVWR--SKYDOPASDINNIYVRMTDAPNEAVLEVAAREAK 175
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 121 GHWGSGVAGVAVSHAKCPVVVVRSDSANED--SKYGVVVGVGDSVSGQATEYAFPAE 179
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 176 LR-----QAPILAGRPPEIREIPDGEFF--RRVODMHHHPDVVRYP 217
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 180 ARGAEIVAVHTWMDMVQASIAAGAAQOOOWDEVERQOTDLTERLAPLVEKYPSTVTKK 239
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 218 ITTHTTGARPLADDERVQGLAVTGG--GEAGQLARLVPSGHPVRHAECSVLVVR 271
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 240 IITDRPVRALAEASNAQLLVGSGHGRGKGMLLASTSRPALQSPAPCPMVRV 294
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
```



```

Db      13 ONIMVAVDGSROSIAFAEALDLADONENHLEFIYSINKLEITHTSAVAFSKTYADEKORT 72
Qy      70 AQSAAVEAAGKLVK-----IETDIPGPAPVPLVEASRDAMI CVSGVIGRYASSI 120
Db      73 EVAMUKKINDAKEXYGAIEHAIVERTGDPRLANV--ICQGENIDIDLIVGATCKAKIQQAL 131
Qy      121 LGSTATLEAKAHCPVAVNR 140
Db      132 VGSTASYVWSHAPCSVLVVK 151

```

RESULT 5

```

US-09-902-540-12843
; Sequence 12843, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 12843
; LENGTH: 430
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-12843

```

```

Query Match      7.4%; Score 103.5; DB 2; Length 430;
Best Local Similarity 24.2%; Pred. No. 0.014;
Matches 78; Conservative 25; Mismatches 118; Indels 101; Gaps 14;

```

```

Qy      11 TIVIGIDGSHAAITPAALMGVDEAISRAPVLRVSVIKPTHPSPDDYDRD-----LAHAER 65
Db      2 SIVCTNTSDAARRASTLAELAKRAGTSIMVYHLP-----DSVAAFGKALGSAEA 55
Qy      66 SLREAQSAVEAAGKLVKTIETDIPRG-PA-----GPVLVEASRDAMI CVSGVIG 114
Db      56 VLSDETKLEQLG--VKVEPVLLTGEPAVLMLEGFCEGASGLVVASRAADESPFGGE--- 110
Qy      115 RYASIIIGSTATLEAKAHCPVAVNR----- 140
Db      111 -----GGTVDRMAQSLTVPLVNRDPAPLEAWVRGERSLKVFLGVDRSLRPFEARDMV 163
Qy      141 -----SKVDQPSADINMIIVVMTDAPDNEAVLEVAAREAKLRQAPILALGSRPELREI 194
Db      164 LTLKSGKAVDVVGAGVYV-----PEEE-----ARRLGKRA--LAFGDALPELQHV 207
Qy      195 PDGEFERRVQDMHHRHPDVRYVPTTHTGIAFLADH-----DERVOLAVIGGEGAOQL 248
Db      208 LRECAELAPLALQGGTPVRA--RVEMGVGR-IADHLVDLAEQSHADLLVGTGTHRRAL 263
Qy      249 ARLVGPSSGHPVRHAECSTLVV 270
Db      264 GRLMSVSRH-ALRLARMSVVCV 284

```

RESULT 6

```

US-09-328-352-5749
; Sequence 5749, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04

```

```

; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 5749
; LENGTH: 1028
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-5749

```

```

Query Match      7.2%; Score 100.5; DB 2; Length 1028;
Best Local Similarity 25.7%; Pred. No. 0.12;
Matches 54; Conservative 33; Mismatches 88; Indels 35; Gaps 12;

```

```

Qy      22 AITAAKGVDEAISRANPLRVSVIKPTHPSPDDYDRDLAAERSLREAQSAVEAAGKLV 81
Db      749 AVSALNMCVNMERR---YKLMSFLKIRKLS--DYNRKV---EELANGEDLLDPTWKPS 800
Qy      82 KIETD--IPR-GPAPVPLVEASRDAMI CVSGVIGRYASSI LGSTATLEAKA-----H 133
Db      801 DSATQERAPRLTLPSPSIYVADERADMI---MOVGKAAEMT-----TRLAKSRAGCIH 852
Qy      134 CPVAVMSKVDQ-----PASDINMIIVVMTDAPDNEAVLEVAAREAKLRQAPILALGQ-- 186
Db      853 LLATQRPSTVDVITGLIKANIPTRVALRVNSKIDSRTILDAGAEIDLGHGDMFLFGPK 912
Qy      187 -RPEELRE--IPDGEFERRVQDMHHR-HPD 212
Db      913 IEPERVHGAFTSDDEVNRICDAMRERGBPD 942

```

RESULT 7

```

US-09-252-991A-27590
; Sequence 27590, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27590
; LENGTH: 542
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27590

```

```

Query Match      7.1%; Score 98.5; DB 2; Length 542;
Best Local Similarity 19.4%; Pred. No. 0.073;
Matches 63; Conservative 45; Mismatches 117; Indels 99; Gaps 13;

```

```

Qy      10 KTIIVGID--GSHAAITPAALMGVDEAISRAPVLR-----SVIKPTHPSPDDYDRDL 60
Db      158 RAIYVIGAKAAGAAVVEKMGQEIQGLVAPFRYGAQCSRIEVEASHRPDDAGERV 217
Qy      61 AHAERSLREAQSAVEAAGKLVKTIETDIPRGAG-----PVLVASRDAL 104
Db      218 A---RVYLELTSGNEDSDRVIFLLSGGSALLSLPAGIGIADQAVNKALKSGAIGE 274
Qy      105 MICV-----GSVIGIRYA-----SSILGSTATLEAB-----KAHCPVAV 138
Db      275 MNCVRKHLISATKGRILAKCPASVFTYAISDVGEDEATVYASPSVGDPTTSDALAIL 334
Qy      139 MRSKVDQPSADINMIIVVMTDAPDNEAVLEVAAREAKLRQAPIL 182
Db      335 RRAIVADPQVNLDWLADPRSETVAPDPCLAHSHFQLIATPQNALDAVAACRAAGSTL 394
Qy      183 ALGSRPEELREIPDGEFERRVQDMHHRHPDVRYVPTTHTGIAFLADHDE--RVQLAVI 240
Db      395 ILGDLBGBAREV-----AKVHGIARQIRKHGQPLAACPVL 431

```

Qy 241 GCGAGQALRVPSGHPFRHAE 264
|||
Db 432 SGGETTIVTRNGRGG---RNAE 451

RESULT 8

US-09-252-991A-30569
; Sequence 30569, Application US/09252991A
; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT FILING DATE: 1999-02-18

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 30569

; LENGTH: 735

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-30569

Query Match 7.1%; Score 98.5; DB 2; Length 735;

Best Local Similarity 24.5%; Pred. No. 0.12; Mismatches 78; Indels 83; Gaps 14;

Matches 63; Conservative 33; Mismatches 78; Indels 83; Gaps 14;

Qy 5 GEPTKTIIVIGIDSHAHTAALMGVDEAISRVAFLVSVIKPTHPSPDDYDRDLAAE 64
|||
Db 297 GAPAG-LVLAQGAELVAG--ALDVLQABEPSRLVGV-----AE 335

Qy 65 RSLBAOS-AVEAAGKLVKLTIDPRGAPRVLEASDAEMICY-GSGVIGRYASSILG 122
|||
Db 336 QVLLAAVALVAAGR---QAEAGSGHGDV---AADAAQAVAGGLGVGGGQAVAD 387

Qy 123 STATELEAKACPVAVMRSKYDOPASDINWIVMTDAPDNEAVLEVAAREAKL---RQ 178
|||
Db 388 AVAADADG---VVGIGVAG---AVVAQAPDQVETIAQAAKQGVYVERQ 434

Qy 179 A-----PILAGGRB-----ELREIPDEFEERRVQDMHNRHPRVVPITHTGI--A 225
|||
Db 435 AGHNLSTGVGGHHRGFAGSVDVRHVADG-----GTGLQVGA 472

Qy 226 RFLADHDERVQAVIYG 242
|||
Db 473 AGLVAVDEQVDGAGYGG 489

Qy 473 AGLVAVDEQVDGAGYGG 489

Qy 473 AGLVAVDEQVDGAGYGG 489

Qy 473 AGLVAVDEQVDGAGYGG 489

Qy 473 AGLVAVDEQVDGAGYGG 489

Qy 473 AGLVAVDEQVDGAGYGG 489

Qy 473 AGLVAVDEQVDGAGYGG 489

Qy 473 AGLVAVDEQVDGAGYGG 489

Qy 473 AGLVAVDEQVDGAGYGG 489

Qy 473 AGLVAVDEQVDGAGYGG 489

Qy 473 AGLVAVDEQVDGAGYGG 489

Qy 473 AGLVAVDEQVDGAGYGG 489

Qy 473 AGLVAVDEQVDGAGYGG 489

Qy 473 AGLVAVDEQVDGAGYGG 489

Qy 473 AGLVAVDEQVDGAGYGG 489

Qy 473 AGLVAVDEQVDGAGYGG 489

Qy 473 AGLVAVDEQVDGAGYGG 489

Qy 473 AGLVAVDEQVDGAGYGG 489

Qy 473 AGLVAVDEQVDGAGYGG 489

Qy 473 AGLVAVDEQVDGAGYGG 489

Qy 473 AGLVAVDEQVDGAGYGG 489

Qy 473 AGLVAVDEQVDGAGYGG 489

Qy 473 AGLVAVDEQVDGAGYGG 489

Qy 473 AGLVAVDEQVDGAGYGG 489

Qy 473 AGLVAVDEQVDGAGYGG 489

Qy 473 AGLVAVDEQVDGAGYGG 489

Qy 473 AGLVAVDEQVDGAGYGG 489

Qy 473 AGLVAVDEQVDGAGYGG 489

Qy 473 AGLVAVDEQVDGAGYGG 489

Qy 473 AGLVAVDEQVDGAGYGG 489

Qy 473 AGLVAVDEQVDGAGYGG 489

Qy 473 AGLVAVDEQVDGAGYGG 489

Qy 473 AGLVAVDEQVDGAGYGG 489

Query Match 7.0%; Score 98; DB 2; Length 657;
Best Local Similarity 22.5%; Pred. No. 0.11;
Matches 71; Conservative 47; Mismatches 108; Indels 90; Gaps 14;

Qy 5 GEPTKTIIVIGIDSHAHTAALMGVDEAISRVAFLVSVIKPTHPSPDDYDRDLAAE 50
|||
Db 291 GEPIMQV-----TSALIRHMAHCHDNDOSQAAALIGIRHLRQLANLGLIKSRRP 344

Qy 51 PSPDDYDRDLAAERSLR-----EASAVEAAGKLVKLTIDPRGAPRVLE 98
|||
Db 345 PAPRAPANNAAGADRELIRIGYGRFSLGILKARQSLTETAPASIGVNVLMSEFPAGQLH 404

Qy 99 ASRAEMICVSGV-----IGRYASSILGSTAE-LAEKACPVAVMRSKYDQ----- 145
|||
Db 405 ALACNE-IDEGTIGEAPVPQAQNSSELMYVAMEPPRAVAVVQGESDIROLSDLRG 463

Qy 146 -----PASDINWIV-----VRMTDAPDNEAVL--EYAREALROAPILALG 185
|||
Db 464 RIALKSNVHMLILQLEADAGLNDVRYVTPPKTPLTASDYLAVDAMMMMDPLSDA 523

Qy 186 GRPELEIRPDGE-----FERRVQDMHNRHPRV--RVVPITHTGIARFLADHDERVQ 236
|||
Db 524 EHTGELRVASGEGRVNHHQFYLRSRDYLAHQGDIMRLLTTLTHG--QFIDSH----- 576

Qy 237 LAVTGGGAGQALRV 252
|||
Db 577 -----RGEAARLL 584

Qy 237 LAVTGGGAGQALRV 252

Qy 237 LAVTGGGAGQALRV 252

Qy 237 LAVTGGGAGQALRV 252

Qy 237 LAVTGGGAGQALRV 252

Qy 237 LAVTGGGAGQALRV 252

Qy 237 LAVTGGGAGQALRV 252

Qy 237 LAVTGGGAGQALRV 252

Qy 237 LAVTGGGAGQALRV 252

Qy 237 LAVTGGGAGQALRV 252

Qy 237 LAVTGGGAGQALRV 252

Qy 237 LAVTGGGAGQALRV 252

Qy 237 LAVTGGGAGQALRV 252

Qy 237 LAVTGGGAGQALRV 252

Qy 237 LAVTGGGAGQALRV 252

Qy 237 LAVTGGGAGQALRV 252

Qy 237 LAVTGGGAGQALRV 252

Qy 237 LAVTGGGAGQALRV 252

Qy 237 LAVTGGGAGQALRV 252

Qy 237 LAVTGGGAGQALRV 252

Qy 237 LAVTGGGAGQALRV 252

Qy 237 LAVTGGGAGQALRV 252

Qy 237 LAVTGGGAGQALRV 252

Qy 237 LAVTGGGAGQALRV 252

Qy 237 LAVTGGGAGQALRV 252

Qy 237 LAVTGGGAGQALRV 252

Qy 237 LAVTGGGAGQALRV 252

Qy 237 LAVTGGGAGQALRV 252

Qy 237 LAVTGGGAGQALRV 252

Qy 237 LAVTGGGAGQALRV 252

Qy 237 LAVTGGGAGQALRV 252

Qy 237 LAVTGGGAGQALRV 252

Qy 237 LAVTGGGAGQALRV 252

Qy 237 LAVTGGGAGQALRV 252

Qy 237 LAVTGGGAGQALRV 252

Qy 237 LAVTGGGAGQALRV 252

Qy 237 LAVTGGGAGQALRV 252

Qy 237 LAVTGGGAGQALRV 252

Qy 237 LAVTGGGAGQALRV 252

Qy 237 LAVTGGGAGQALRV 252

Qy 237 LAVTGGGAGQALRV 252

Qy 237 LAVTGGGAGQALRV 252

Qy 237 LAVTGGGAGQALRV 252

Qy 237 LAVTGGGAGQALRV 252

Qy 237 LAVTGGGAGQALRV 252

Qy 237 LAVTGGGAGQALRV 252

Qy 237 LAVTGGGAGQALRV 252

Qy 237 LAVTGGGAGQALRV 252

Qy 237 LAVTGGGAGQALRV 252

Qy 237 LAVTGGGAGQALRV 252

Qy 237 LAVTGGGAGQALRV 252

Qy 237 LAVTGGGAGQALRV 252

Qy 237 LAVTGGGAGQALRV 252

Qy 237 LAVTGGGAGQALRV 252

Qy 237 LAVTGGGAGQALRV 252

Qy 237 LAVTGGGAGQALRV 252

Db 645 ARHGD-----LVIYIGIQAQWGLEDKVFGIQRRIITDAPGSLIV 685

RESULT 11

US-09-107-433-3293

Sequence 3293, Application US/09107433

Patent No. 6800744

GENERAL INFORMATION:

APPLICANT: Lynn A Doucette-Stamm and David Bush

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE FOR DIAGN

SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE

THERAPEUTICS

NUMBER OF SEQUENCES: 5206

CORRESPONDENCE ADDRESS:

ADDRESSEE: GENOME THERAPEUTICS CORPORATION

STREET: 100 Beaver Street

CITY: Waltham

STATE: Massachusetts

COUNTRY: USA

ZIP: 02354

COMPUTER READABLE FORM:

MEDIUM TYPE: CD-ROM ISO9660

OPERATING SYSTEM: <Unknown>

SOFTWARE: <Unknown>

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/107,433

FILING DATE: 30-Jun-1998

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/085131

FILING DATE: May 12, 1998

APPLICATION NUMBER: 60/051553

FILING DATE: July 2, 1997

ATTORNEY/AGENT INFORMATION:

NAME: Ariniello, Pamela Deneke

REGISTRATION NUMBER: 40,489

REFERENCE/DOCKET NUMBER: GTC-011

TELECOMMUNICATION INFORMATION:

TELEPHONE: (781)893-5007

TELEFAX: (781)893-8277

INFORMATION FOR SEQ ID NO: 3293:

SEQUENCE CHARACTERISTICS:

LENGTH: 148 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: YES

ORIGINAL SOURCE:

ORGANISM: Streptococcus pneumoniae

FEATURE:

NAME/KEY: misc feature

LOCATION: (B) LOCATION 1...148

SEQUENCE DESCRIPTION: SEQ ID NO: 3293:

US-09-107-433-3293

Query Match 6.9%; Score 97; DB 2; Length 148;
Best Local Similarity 25.4%; Pred. No. 0.014;
Matches 36; Conservative 27; Mismatches 69; Indels 10; Gaps 4;

10 KTIIVIGDSHAATITAAIMGVDAISRAVPLRLVSVIKP-THPSDDYDRDL-----AHA 63

4 QNIMVAIDGSKEDDLAFVKGVSHALRNDAKLTIAHVIDTRALQSVSTFDAYEVELQYDA 63

64 ERSIREAQSVAEAGKLVKIEITDIPRGAPVIVASRDAB---MICVSGVIGRYVASSI 120

64 ESLMKEYEKAKXAG-VADVHIVIEEMGNPKTLARTIPDAEVDLLVGATGLNAFERLL 122

121 LGSTATLEAKAHCPVAVWRSK 142

123 VGSSEYILRHAKVDLLVVRQ 144

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

US-09-583-110-3588

Sequence 3588, Application US/09583110

Patent No. 6699703

GENERAL INFORMATION:

APPLICANT: Lynn Doucette-Stamm et al.

TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus

FILE REFERENCE: PATH00-07A

CURRENT APPLICATION NUMBER: US/09/583,110

CURRENT FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: US 09/107,433

PRIOR FILING DATE: 1998-06-30

PRIOR APPLICATION NUMBER: US 60/085,131

PRIOR FILING DATE: 1998-05-12

PRIOR APPLICATION NUMBER: US 60/051,553

PRIOR FILING DATE: 1997-07-02

NUMBER OF SEQ ID NOS: 5322

SEQ ID NO 3588

LENGTH: 150

TYPE: PRT

ORGANISM: Streptococcus pneumoniae

US-09-583-110-3588

Query Match 6.9%; Score 97; DB 2; Length 150;
Best Local Similarity 25.4%; Pred. No. 0.014;
Matches 36; Conservative 27; Mismatches 69; Indels 10; Gaps 4;

10 KTIIVIGDSHAATITAAIMGVDAISRAVPLRLVSVIKP-THPSDDYDRDL-----AHA 63

6 QNIMVAIDGSKEDDLAFVKGVSHALRNDAKLTIAHVIDTRALQSVSTFDAYEVELQYDA 65

64 ERSIREAQSVAEAGKLVKIEITDIPRGAPVIVASRDAB---MICVSGVIGRYVASSI 120

64 ESLMKEYEKAKXAG-VADVHIVIEEMGNPKTLARTIPDAEVDLLVGATGLNAFERLL 124

121 LGSTATLEAKAHCPVAVWRSK 142

125 VGSSEYILRHAKVDLLVVRQ 146

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

TELECOMMUNICATION INFORMATION:
TELEPHONE: (781) 893-5007
TELEFAX: (781) 893-8277
INFORMATION FOR SEQ ID NO: 3842:
SEQUENCE CHARACTERISTICS:
LENGTH: 161 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEICAL: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1..161
SEQUENCE DESCRIPTION: SEQ ID NO: 3842:
US-09-107-532A-3842

Query Match 6.9%; Score 97; DB 2; Length 161;
Best Local Similarity 25.2%; Pred. No. 0.016;
Matches 38; Conservative 30; Mismatches 71; Indels 12; Gaps 5;

QY 10 KTIIVIGDSHAATTALMGVDEAISRVPRLVSVIKP--THPSDDYDRDLA-----A 63
DB 8 KKIIVAVDSDSEALAFKAVVAIRNNGELVLAHVITRTSFQVSSFDGMLAEQATBMA 67
QY 64 ERSIREAQSVAEAGKLVKIEIDI PRGPAPVL---VEASRDAMICVSGVIGRYASSI 120
DB 68 KQTLADVESNKKAG--LNNVTSVVEYSGPKQIIAREIPEDNQVDLIMGATGLNAVERLF 126
QY 121 LGSTATLAEAKHCPCVAVMSKVDQ--PASD 149
DB 127 IGSVSEYVIRNAACDVAVRTDLENOLPAKD 157
RESULT 14
US-09-252-991A-30428
; Sequence 30428, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 30428
; LENGTH: 420
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30428

Query Match 6.9%; Score 97; DB 2; Length 420;
Best Local Similarity 25.4%; Pred. No. 0.072;
Matches 58; Conservative 24; Mismatches 74; Indels 72; Gaps 13;

QY 51 PSPD-DYDR-DLAHRSRLREAGSVAAGKLVKIEIDI PRGPA-----GPVLVEASR 101
DB 216 PAPERNDRLPRLNVAHL-DAQARQVFLG-----DSPGHAHGGIARRGTAAVA 267
QY 102 DAEMICVSGVIGRYASSIIGSTATLAEAKHCPCVAVMSKVDQPASDINMIYVMTDAP 161
DB 268 HAVLVVGVVGMGR-AEQVLDRRVV-----LGLLVGALDQA-----DRR 306
QY 162 DNEAVLEVARBEAKLRQATLALGAP-----EELREIPDGEFERRVQDMHHRPDV 213
DB 307 AGGLAHEPGK--NLHLVGFALGQVPAAGLAFVEVALQVPGGQFQPRGA----- 355

QY 214 RVYPITHTGIARFLADHDERVOLAVYGGEGQALRLVPSGHPVR 261
DB 356 ---FID-----DGDORRAVAFACGDBGELA--VGVAGHARLR 388

RESULT 15
US-09-328-352-7302
; Sequence 7302, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 7302
; LENGTH: 157
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-7302

Query Match 6.7%; Score 94; DB 2; Length 157;
Best Local Similarity 25.4%; Pred. No. 0.033;
Matches 36; Conservative 27; Mismatches 63; Indels 16; Gaps 5;

QY 12 IIVIGDS-----HAITTAALMGVDEAISRVP--PLRLVSVIKPTHSPDDYDRDL 60
DB 16 ILVPVDSPTSLIANNHAAFLAKFSKTVLVVALTIIDPISVEIIDSITELADYENKAR 75
QY 61 AHAERSLREAGSVAEAGKLVKIEIDI PRGPA--GPVLVEASRD--AEVICVSGVIGRYA 117
DB 76 ASIGSILDOAKQEQFSOHG--ISVETKIVGGQTIHTEIKATLAKADLLVIGSHGKGRK 133
QY 118 SSILGSTATLAEAKHCPCVAVM 139
DB 134 KPFLGSVTQALLGEIHVPVLV 155

Search completed: March 23, 2006, 05:30:16
Job time : 19.6513 secs


```

US-10-617-038-21
; Sequence 21, Application US/10617038
; Publication No. US20040057963A1
; GENERAL INFORMATION:
; APPLICANT: Andersen, Peter
; APPLICANT: Rosenkrands, Ida
; APPLICANT: Stryhn, Anette
; TITLE OF INVENTION: Therapeutic TB Vaccine
; FILE REFERENCE: S515AUSA
; CURRENT APPLICATION NUMBER: US/10/617,038
; CURRENT FILING DATE: 2003-07-11
; PRIOR APPLICATION NUMBER: DK PA 2002 01098
; PRIOR FILING DATE: 2002-07-13
; PRIOR APPLICATION NUMBER: US 60/401,725
; PRIOR FILING DATE: 2002-08-07
; NUMBER OF SEQ ID NOS: 187
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 21
; LENGTH: 279
; TYPE: PRF
; ORGANISM: Mycobacterium tuberculosis
US-10-617-038-21

```

[illegible]

```

RESULT 3
US-10-617-038-42
; Sequence 42, Application US/10617038
; Publication No. US20040057963A1
; GENERAL INFORMATION:
; APPLICANT: Andersen, Peter
; APPLICANT: Rosenkrands, Ida
; APPLICANT: Stryhn, Anette
; TITLE OF INVENTION: Therapeutic TB Vaccine
; FILE REFERENCE: SSISADUSA
; CURRENT APPLICATION NUMBER: US/10/617,038
; CURRENT FILING DATE: 2003-07-11
; PRIOR APPLICATION NUMBER: DK PA 2002 01098
; PRIOR FILING DATE: 2002-07-13
; PRIOR APPLICATION NUMBER: US 60/401,725
; PRIOR FILING DATE: 2002-08-07
; NUMBER OF SEQ ID NOS: 187
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 42
; LENGTH: 268
; TYPE: PR1
; ORGANISM: Mycobacterium tuberculosis
US-10-617-038-42
Query Match      25.6%; Score 357.5; DB 4; Length 268

```

[illegible]

```

RESULT 4
US-10-138-473-34
Sequence 34, Application US/10138473
Publication No. US20030165525X1
GENERAL INFORMATION:
APPLICANT: ANDERSEN, Peter
APPLICANT: WELDINGH, Karln
APPLICANT: HANSEN, Christina Veggerby
APPLICANT: FLORIO, Walter
APPLICANT: OKKELS, Li Mei Meng
APPLICANT: SKJORT, Rikke Louise Vinther
APPLICANT: ROSENKRANDS, Ida
APPLICANT: RASMUSSEN, Peter Birk
TITLE OF INVENTION: TB Diagnostic Based On Antigens From M. tuberculosis
FILE REFERENCE: 0459-0710P
CURRENT APPLICATION NUMBER: US/10/138,473
CURRENT FILING DATE: 2002-05-02
PRIORITY APPLICATION NUMBER: US 10/060,428
PRIORITY FILING DATE: 2002-01-29
PRIORITY APPLICATION NUMBER: US 09/415,884
PRIORITY FILING DATE: 1999-10-08
PRIORITY APPLICATION NUMBER: US 60/116,673
PRIORITY FILING DATE: 1999-01-21
PRIORITY APPLICATION NUMBER: DK 1998 01281
PRIORITY FILING DATE: 1998-10-18
PRIORITY APPLICATION NUMBER: US 60/070,488
PRIORITY FILING DATE: 1998-01-05
PRIORITY APPLICATION NUMBER: DK 1997 01277
PRIORITY FILING DATE: 1997-11-10
PRIORITY APPLICATION NUMBER: US 60/044,624
PRIORITY FILING DATE: 1997-04-18
PRIORITY APPLICATION NUMBER: DK 1997 00376
PRIORITY FILING DATE: 1997-04-02
NUMBER OF SEQ ID NOS: 174
SOFTWARE: PatentIn version 3.1
SEQ ID NO 34
LENGTH: 297
TYPE: PR1
ORGANISM: Mycobacterium tuberculosis
US-10-138-473-34

```

Query Match 19.1%; Score 267; DB 4; Length 297;
 Best Local Similarity 29.3%; Pred. No. 6,66-17;
 Matches 64; Conservative 42; Mismatches 131; Indels 30; Gaps 8

12 IYVGDSSHAITATALMGVDEALSRAPVLRVSYIKP-----THPS-----DYYDR 58
 |||||
 10 IYVGDIDSPAAQVAVRWAAADDAELRKLPLTLTAHVSEVATWLEVPPLPGVLRMQODHGR 69
 |||||


```
Qy 59 DLAAERSLREAQSAVEAAGKLVKIEITDIPRPAQPVLYEASRDAMI CVSGVGIGRYAS 118
Db 70 HL--IDDAIKVVEQASLFAAGP-PTVHSEIVPAAAVPTLVDMSKDAVIMVVGCLSGRMPG 126
Qy 119 SLIGSTATELAEKAKHCPVAVRSKYD-QPASDINMIVVMTDADNEAVLEYAAREAKLR 177
Db 127 RLSSVSSGLRHHACPVVITIHDEDSVPHPOQAPVLGVGSSASSELATIADEASRR 186
Qy 178 QAPITALGGRPE-ELREIP-----DGEFERRVQDWHHHPDVRVYPITTHGIA 225
Db 187 NVDLVALHMSDVVSEWPGIDWPATQSMAGVLAERLAGQGERYPNVAITRVVVRDOPA 246
Qy 226 RFLADHDERVOLAVIGG-GEAGQLARLVGSGHPVFRHAECSVLVVR 271
Db 247 RQLVGRSEAOVLVVVSGRGGYAGMLVGSGETVAQIARFPVIVAR 293
```

RESULT 5
US-10-450-726-4
Sequence 4, Application US/10450726
Publication No. US20040242471A1
GENERAL INFORMATION:
APPLICANT: Dick, Thomas
TITLE OF INVENTION: DORMANCY-INDUCED MYCOBACTERIUM PROTEINS
FILE REFERENCE: 50318/002001
CURRENT APPLICATION NUMBER: US/10/450,726
CURRENT FILING DATE: 2003-06-13
PRIOR APPLICATION NUMBER: PCT/EP01/14551
PRIOR FILING DATE: 2001-12-11
PRIOR APPLICATION NUMBER: 0030368.5
PRIOR FILING DATE: 2000-12-13
NUMBER OF SEQ ID NOS: 21
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4
LENGTH: 297
TYPE: PRT
ORGANISM: Mycobacterium tuberculosis
US-10-450-726-4

```
Query Match 19.1%; Score 267; DB 5; Length 297;
Best Local Similarity 29.3%; Pred. No. 6.6e-17;
Matches 84; Conservative 42; Mismatches 131; Indels 30; Gaps 8;

Qy 12 IIVGIDGSHAATTAALMGVDEAISRAVPLRLVSYIKP-----THPSP-----DDYDR 58
Db 10 IIVGIDGSPAQAQVAVRAAARDAELRKIFLITVHAVSEPVATWLEVPRLPGVLRMQDHR 69
Qy 59 DLAAERSLREAQSAVEAAGKLVKIEITDIPRPAQPVLYEASRDAMI CVSGVGIGRYAS 118
Db 70 HL--IDDAIKVVEQASLFAAGP-PTVHSEIVPAAAVPTLVDMSKDAVIMVVGCLSGRMPG 126
Qy 119 SLIGSTATELAEKAKHCPVAVRSKYD-QPASDINMIVVMTDADNEAVLEYAAREAKLR 177
Db 127 RLSSVSSGLRHHACPVVITIHDEDSVPHPOQAPVLGVGSSASSELATIADEASRR 186
Qy 178 QAPITALGGRPE-ELREIP-----DGEFERRVQDWHHHPDVRVYPITTHGIA 225
Db 187 NVDLVALHMSDVVSEWPGIDWPATQSMAGVLAERLAGQGERYPNVAITRVVVRDOPA 246
Qy 226 RFLADHDERVOLAVIGG-GEAGQLARLVGSGHPVFRHAECSVLVVR 271
Db 247 RQLVGRSEAOVLVVVSGRGGYAGMLVGSGETVAQIARFPVIVAR 293
```

RESULT 6
US-09-738-626-6745
Sequence 6745, Application US/09738626
Publication No. US20020197605A1
GENERAL INFORMATION:
APPLICANT: NAKAGAWA, SATOSHI
APPLICANT: MIZOGUCHI, HIROSHI
APPLICANT: ANDO, SEIKO

```
APPLICANT: HAYASHI, MIKIRO
APPLICANT: OCHIAI, KEIKO
APPLICANT: YOKOI, HARUHIKO
APPLICANT: TATEISHI, NAKO
APPLICANT: SENOH, AKIHIRO
APPLICANT: IKEDA, MASATO
APPLICANT: OZAKI, AKIO
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-125
CURRENT APPLICATION NUMBER: US/09/738,626
CURRENT FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: JP 99/377484
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: JP 00/159162
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: JP 00/280988
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 7059
SOFTWARE: Patentin ver. 3.0
SEQ ID NO 6745
LENGTH: 301
TYPE: PRT
ORGANISM: Corynebacterium glutamicum
US-09-738-626-6745
```

```
Query Match 19.0%; Score 266; DB 3; Length 301;
Best Local Similarity 28.8%; Pred. No. 8.4e-17;
Matches 85; Conservative 45; Mismatches 131; Indels 34; Gaps 8;

Qy 8 TMTIIVGIDGSHAATTAALMGVDEAISRAVPLRLVSYIKP-----PDYDR 58
Db 3 TEDIVAAVDSQASAKAVMAANTANKRIGIPRLAS--SYTFQFLYAGWVPPQSLFD 60
Qy 59 DLAAERSLREAQSAVEAAGKLVKIEITDIPRPAQPVLYEASRDAMI CVSGVGIGRYA 117
Db 61 DLQAEALKEKINEARDIAHEPAPEIKGHTTAESPIDMLEMSPDAMITMGSRGGLS 120
Qy 118 SSLIGSTATELAEKAKHCPVAVMR--SKVDQASDINMIVVMTDADNEAVLEYAAREAK 175
Db 121 GWMVSVSAGVAVSHAKCPVVVREDSAVNED-SKYGVVVGVGSEVSQATEYAFEALE 179
Qy 176 LR-----QAPILAGRPPELRIPDGEFE---RRVODMHHRRPDVRYP 217
Db 180 ARGAEIVAVHTWMDQVQSLAGIAAQQOWDEVERQOTDMLERLAPVEKXPSTVVK 239
Qy 218 ITTHGTARFLADHDERVOLAVIGG-GEAGQLARLVGSGHPVFRHAECSVLVVR 271
Db 240 IITRDPRVALAASENAQLVVVSGHGGFKGMLGSTRALLQSAFCPMVVR 294
```

RESULT 7
US-09-746-660A-6
Sequence 6, Application US/09746660A
Publication No. US2003049804A1
GENERAL INFORMATION:
APPLICANT: Pompejus, Markus
APPLICANT: Krogger, Burthard
APPLICANT: Schroder, Hartwig
APPLICANT: Zelder, Oskar
APPLICANT: Habermayer, Gregor
APPLICANT: Kim, Jun-Won
APPLICANT: Lee, Heung-Schick
APPLICANT: Hwang, Byung-Joon
TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING METABOLIC PATHWAY PROTEINS
FILE REFERENCE: BGI-121CP2
CURRENT APPLICATION NUMBER: US/09/746,660A
CURRENT FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 09/606740
PRIOR FILING DATE: 2000-06-23
PRIOR APPLICATION NUMBER: 09/603124
PRIOR FILING DATE: 2000-06-23
PRIOR APPLICATION NUMBER: 60/141031

Db 240 IITDRPVRALAEASENAOQLVVGSHGKFGKMGILGSTSRALLQSAFCPMVVVR 294

RESULT 10
US-11-006-098-314

```

; Sequence 314, Application US/11006098
; Publication No. US20050153402A1
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schröder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauser, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING
; TITLE OF INVENTION: REGULATORY PROTEINS
; FILE REFERENCE: BGI-123CP
; CURRENT APPLICATION NUMBER: US/11/006,098
; CURRENT FILING DATE: 2004-12-06
; PRIOR APPLICATION NUMBER: US/09/602,874
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 60/142690
; PRIOR FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: 60/151251
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19930476.9
; PRIOR FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: DE 19931419.5
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: DE 19931420.9
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19932122.1
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932128
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932206.6
; PRIOR FILING DATE: 1999-07-09
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 362
; SEQ ID NO 314
; LENGTH: 301
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-11-006-098-314

```

Query Match 19.0%; Score 266; DB 6; Length 301;
Best Local Similarity 28.8%; Pred. No. 8.4e-17;
Matches 85; Conservative 45; Mismatches 131; Indels 34; Gaps 8;

```

QY 8 TMTIIVIGIDGSHAAITPAALMGVDEAISRAVPLRLVSVIKFTHS-----PDDYDR 58
DB 3 TEDIVVAIVDGSDDSKQAVRWAAANTANKRGIPRLIAS--SYTMPQFLYAEQWVPPQELFD 60
QY 59 DL-AHARSRSREASAVEAAGKLVKTIENDIPRGAPVLYEASRDAMI CVSGVIGIRYA 117
DB 61 DLQAEALAEKINEADIAHEVAPEIKIGTTIAGSPIDMLMSDPATYITVWGSRGIGLS 120
QY 118 SSIIIGSTATLEAKHCPVAVMR--SKVDQPSADINMIVAMTDA PDNEAVLEVAABEAK 175
DB 121 GMMVGSVGAIVSHAKCPVVVVRSDSAVNEB-SKKGPPVVGVDSDSVSQATEYAFABAE 179
QY 176 LR-----CAPILALGGRPEELREIPDGEF--RRVQDMHHRHDPVRYVP 217
DB 180 ARGAEIIVAHVTMDMOVQASLAGLAAQOQWDEVERQOTDMLIRLAPLVYKESPVYVKK 239
QY 218 ITHHTGATPLADDERVQOLAVIGG-GEAGQLARLVGSPGHVFRHAECSVLYVR 271
DB 240 IITDRPVRALAEASENAOQLVVGSHGKFGKMGILGSTSRALLQSAFCPMVVVR 294

```

RESULT 11

US-10-282-122A-53861
Sequence 53861, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:

```

; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Foreysch, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 53861
; LENGTH: 300
; TYPE: PRT
; ORGANISM: Corynebacterium diptheriae
US-10-282-122A-53861

```

Query Match 17.8%; Score 249; DB 4; Length 300;
Best Local Similarity 28.5%; Pred. No. 3.6e-15;
Matches 84; Conservative 46; Mismatches 123; Indels 42; Gaps 10;

```

QY 12 IIVGIDGSHAAITPAALMGVDEAISRAVPLRLVSVIKFTHSPPD-YDRDLAHAERSIREA 70
DB 7 VVAIVDSEASQAVRWAAANTANKRGVPLRLIAA-----SYTMPQFLYAEQWVPPQELFD 62
QY 71 QS-----AVEAAGKL-----VKTIENDIPRGAPVLYEASRDAMI CVSGVIGIRYASS 119
DB 63 QSETMDIIEARVAHEVADIKIGYVIAEGSPIDMLDMSDVYTIMVWGSRGIGLSGM 122
QY 120 ILGSTATLEAKHCPVAVMRSKVDQPSADINM---IIVAMTDA PDNEAVLEVAABEAKL 176
DB 123 VMGSVSAVVSADCPVVVVR--DHNVTETNKKGPVVGVDSDSVSQATEYAFABEAOA 180
QY 177 R-----CAPILALGGRPEELREIPDGE---FRRVQDMHHRHDPVRYVP 217
DB 181 RGAKLVAIHTMDMOVQASLAGLAAQOEW-EIIEKEQFTLLKDRLOPLERFPDVEVEM 239
QY 218 ITHHTGATPLADDERVQOLAVIGG-GEAGQLARLVGSPGHVFRHAECSVLYVR 271
DB 240 VITDRPVRALAECAHNAOQLVVGSHGKFGKMGILGSTSRALLQSAFCPMVVVR 294

```

```

RESULT 12
US-10-156-761-8814
; Sequence 8814, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 8814
; LENGTH: 290
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-8814

```

```

Query Match 17.4%; Score 242.5; DB 4; Length 290;
Best Local Similarity 29.7%; Pred. No. 1.4e-14;
Matches 87; Conservative 38; Mismatches 123; Indels 45; Gaps 11;

```

```

QY 12 IIIVGIDGSHAITAALGVDEAISRVPRLVSVIKPTHPSPDDYDRLHAERSLRE-- 69
   :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 5 LVGVDSSESMRAVDMAADALRGVPLRV-----YASLWRYGSAALADGLGRPE 58
   :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 70 ---AQSVAEAGK-----VKETDI-PRGPGPVLEASRDAMICVSGVIGRYAS 118
   :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 59 QVIAADDIVEAAQAHRRHADVKISTDVLPEEPA-PALLRGNAFALVIGSRSGIAE 117
   :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 119 SIIGSTAELAEKACPVAVNRSKYDQ--PASDINMIVVMTDAPDNEAVLEVAAREKL 176
   :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 118 LLLGSVSLAVAAKAYCPVITVIRGSDNRAGCGRRIVLVGSEPPDPAVRFVAEARA 177
   :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 177 ROAPILALGRPEELRE-IPDEFEERYVDWHNRH-----PDVRYVPIITHT 222
   :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 178 RGCVAEAVRAMRRARERALEGHLPSRRAARLHERATEVLEAVLSGAPPDVELRRAVEG 237
   :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 223 GIARFLADHDERVOLAVIYG---GEAG-QLARLVGSGEPVFRHAECSVLVV 270
   :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 238 PARKVLLDASATADLLVVGARHHPGHSGLQIGRV----AHAMLHHSACPVAVV 286
   :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

```

```

RESULT 13
US-10-156-761-8839
; Sequence 8839, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02

```

```

; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 8839
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-8839

```

```

Query Match 15.0%; Score 210; DB 4; Length 295;
Best Local Similarity 29.0%; Pred. No. 2e-11;
Matches 87; Conservative 37; Mismatches 124; Indels 52; Gaps 12;

```

```

QY 10 KTIIVGIDGSHAITAALGVDEAISRVPRLVSVIKPTHPSP--DDVD---RLDAA 63
   :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 6 RTTVGIDGSRGTGLDADWAAERARRRLRLHGAEPVAPVDDVLPARTALDRA 65
   :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 64 ERLREASVAEAGKLVKIETDIPRGAPGVLEASRDAMICVSGVIGRYASSILGS 123
   :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 66 AVQLAVAHPLDIIAR-RLET-----PAVPLALAAARETLVLSRGFTGAGFLVGS 118
   :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 124 TATELAEKACPVAVNRS---KVDQASD-----INMIVVMTDAPDNEAVLEY 169
   :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 119 VALAVAAARAPVVLVTRGELAEDEHMPADDGTPGRAAVLPVGLDLDPH-ADRLIAY 177
   :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 170 AAREAKLRQAP-----LALGRPEELREIPDGEFERRVQDMHNRHP- 211
   :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 178 AFBAAVRSAPLHVHTWTTLPPTGAPGAPLPDGAAREEGARRLTTLTQPMRHKFPG 237
   :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 212 -DVRVYPIITHTGIARFLADHDERVOLAVIGGEAGQLARLVGSGHPVFRHAECSVLVV 270
   :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 238 TDVLERVIPGHPG--HHLRASIRAGLVIGRTS---APGLQRAARSLIHAGCPVAVV 292
   :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

```

```

RESULT 14
US-10-437-963-153439
; Sequence 153439, Application US/10437963
; Publication No. US2004012343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 153439
; LENGTH: 185
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_53395C.1.pep
US-10-437-963-153439

```

```

Query Match 10.3%; Score 143.5; DB 4; Length 185;
Best Local Similarity 28.7%; Pred. No. 2.5e-05;
Matches 48; Conservative 25; Mismatches 59; Indels 35; Gaps 7;

```

```

QY 5 GEP-----IMKTIIVGIDGSHAITAALGVDEAISRVPRLVSVIKPTHPSPDDYDRD 59
   :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 16 GEGGAAAGTMR-VVAVNDASESINALSMALDNVIGRRAGAVSVVYVHAQH-GPDHFPV 73
   :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 60 L-----AAERSLREASVAEAGKLVKIETD-----IPRGAPGV 95
   :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 74 VAHIGISIAVAPASIESMRKQOE--EISRKVVSALDVCKOREVATGALVEGDAKKA 131
   :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 96 LVEASD--AEMICVSGVIGRYASSITGSTATLEAKHCHCPVAVNR 140
   :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

```


This page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocelelation Ltd.

OM protein - protein search, using sw model

Run on: March 23, 2006, 06:17:23 ; Search time 6.52731 Seconds
(without alignments)
1192.746 Million cell updates/sec

Title: US-10-617-038-27

Perfect score: 1397

Sequence: 1 MSGGEPMTKTIIVGIDSH.....GPSGHPVFRHAECSVLVRR 272

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 169630 seqs, 2862289 residues

Total number of hits satisfying chosen parameters: 169630

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA_New:*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US05_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	266	19.0	301	US-11-055-822-14	Sequence 14, Appl
2	262.5	18.8	305	US-11-087-099-3468	Sequence 3468, Ap
3	149	10.7	152	US-11-087-099-6347	Sequence 6347, Ap
4	130.5	9.3	163	US-11-087-099-11409	Sequence 11409, A
5	130.5	9.3	163	US-11-087-099-11409	Sequence 11409, A
6	130.5	9.3	163	US-11-087-099-11409	Sequence 11409, A
7	128.5	9.2	171	US-11-087-099-505	Sequence 4790, Ap
8	128.5	9.2	171	US-11-087-099-505	Sequence 505, App
9	125.5	9.0	153	US-11-096-568A-15853	Sequence 15853, A
10	122.5	8.8	174	US-11-096-568A-27471	Sequence 27471, A
11	122	8.7	171	US-11-096-568A-32572	Sequence 32572, A
12	122	8.7	179	US-11-096-568A-27882	Sequence 27882, A
13	120.5	8.6	151	US-11-096-568A-15855	Sequence 15855, A
14	120	8.6	172	US-11-087-099-11256	Sequence 11256, A
15	118.5	8.5	175	US-11-087-099-7223	Sequence 7223, Ap
16	117.5	8.4	152	US-11-096-568A-4792	Sequence 4792, Ap
17	116.5	8.3	169	US-11-087-099-436	Sequence 436, App
18	116.5	8.3	169	US-11-087-099-436	Sequence 436, App
19	116	8.3	169	US-11-087-099-436	Sequence 436, App
20	114	8.2	169	US-11-087-099-4994	Sequence 4994, Ap
21	112	8.0	182	US-11-087-099-2660	Sequence 2660, Ap
22	109	7.8	186	US-11-087-099-7765	Sequence 7765, Ap
23	107.5	7.7	150	US-11-087-099-3411	Sequence 3411, Ap
24	107	7.7	175	US-11-087-099-9020	Sequence 9020, Ap
25	106.5	7.6	162	US-11-096-568A-1508	Sequence 1508, Ap

26	106.5	7.6	181	US-11-087-099-10726	Sequence 10726, A
27	103.5	7.4	162	US-11-087-099-5731	Sequence 5731, Ap
28	103	7.4	159	US-11-087-099-12270	Sequence 12270, A
29	103	7.4	161	US-11-087-099-2546	Sequence 2546, Ap
30	101.5	7.3	162	US-11-087-099-9120	Sequence 9120, Ap
31	101.5	7.3	162	US-11-087-099-10134	Sequence 10134, A
32	101.5	7.3	203	US-11-096-568A-10261	Sequence 10261, A
33	101	7.2	153	US-11-096-568A-1509	Sequence 1509, Ap
34	100.5	7.2	160	US-11-087-099-3178	Sequence 3178, Ap
35	99	7.1	125	US-11-087-099-6096	Sequence 6096, Ap
36	99	7.1	161	US-11-087-099-6989	Sequence 6989, Ap
37	99	7.1	162	US-11-096-568A-10262	Sequence 10262, A
38	99	7.1	171	US-11-087-099-8466	Sequence 8466, Ap
39	99	7.1	171	US-11-087-099-8499	Sequence 8499, Ap
40	98.5	7.1	270	US-10-467-657-1630	Sequence 1630, Ap
41	98	7.0	171	US-11-087-099-7706	Sequence 7706, Ap
42	97.5	7.0	164	US-11-087-099-6603	Sequence 6603, Ap
43	97.5	7.0	3655	US-11-075-185-5	Sequence 5, Appl
44	97	6.9	154	US-11-096-568A-27472	Sequence 27472, A
45	97	6.9	179	US-11-096-568A-6711	Sequence 6711, Ap

ALIGNMENTS

RESULT 1
US-11-055-822-14
Sequence 14, Application US/11055822
Publication NO. US20050260707A1
GENERAL INFORMATION:
APPLICANT: Pompeius, Markus
APPLICANT: Kroege, Burkhard
APPLICANT: Schröder, Hartwig
APPLICANT: Zeldner, Oskar
APPLICANT: Haberhauser, Gregor
TITLE OF INVENTION: CORINNEBACTERIUM GLUTAMICUM GENES ENCODING
FILE REFERENCE: BGI-121CPCN
CURRENT APPLICATION NUMBER: US/11/055, 822
CURRENT FILING DATE: 2005-02-11
PRIOR APPLICATION NUMBER: 09/606,740
PRIOR FILING DATE: 2000-06-23
PRIOR APPLICATION NUMBER: 60/141,031
PRIOR FILING DATE: 1999-06-25
PRIOR APPLICATION NUMBER: 60/142,101
PRIOR FILING DATE: 1999-07-02
PRIOR APPLICATION NUMBER: 60/148,613
PRIOR FILING DATE: 1999-08-12
PRIOR APPLICATION NUMBER: 60/187,970
PRIOR FILING DATE: 2000-03-09
PRIOR APPLICATION NUMBER: DE 19930476.9
PRIOR FILING DATE: 1999-07-01
PRIOR APPLICATION NUMBER: DE 19931415.2
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931418.7
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931419.5
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931420.9
PRIOR FILING DATE: 1999-07-08
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 1158
SEQ ID NO 14
LENGTH: 301
TYPE: PRT
ORGANISM: Corynebacterium glutamicum
US-11-055-822-14

Query Match 19.0%; Score 266; DB 7; Length 301;
Best Local Similarity 28.8%; Pred. No. 1,1e-15;
Matches 85; Conservative 45; Mismatches 131; Indels 34; Gaps 8;
Qy 8 TMKTIIVGIDSHAATPALMGVDEAISRAVPLRLVSVIKRTHPS-----PDDYDR 58

```

Db      3 TEDIIVAAVDDSDSKQAVRRAAATANKRGIPRLAAS--SYTMQFLYAEQMPVQBEFD 60
Qy      59 DL-AHAERSLREASAVEAAGKLVKLETIDIPRGAPVLYBASRDAEMICVSGVIGRYA 117
        61 DLQAELEKITEADIDIAHEVAPEIKGHTIAEGSPIDMLMSPDATIMVSGSLGLGSL 120
Qy      118 SSLIGSTATELAERKACPAVAVMR--SKYDOPASDINMIIVMTDAPDNEAVLEAYAREAK 175
        121 GMVWGSVSGAVVSHAKCPVVVVVREDSAVNED--SKYGPVVGVDSGEVSGQATEYAFABAE 179
Db      176 LR-----QAPILAGRPELEIRPDGEF--RRVQMHHRHDPVAVYR 217
Qy      180 ARGAEIIVATMTDMQVQASLAGLAAQQQWDEVERQDTMLIRLAPLYEKYPSVTVYK 239
Db      218 ITTHTGIAFLADDERVOLAVIGG-GEAGQIARLVGPSGHPVRHAECSVLVVR 271
        240 IITDRPRALAESENQQLLVVSGHGRGFKGMLLSTSLALQSAFCPMVVR 294

```

```

RESULT 2
US-11-087-099-3468
; Sequence 3468, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 3468
; LENGTH: 305
; TYPE: PRT
; ORGANISM: Thermobifida fusca
US-11-087-099-3468

```

```

Query Match      18.8%; Score 262.5; DB 7; Length 305;
Best Local Similarity 28.7%; Pred. No. 2,3e-15;
Matches 89; Conservative 41; Mismatches 131; Indels 49; Gaps 8;

```

```

Qy      1 MSGGGEPTMTIIVIGIDGSHAITTAALMGVDEAISRAVPLRLV-----SYTKP----- 48
        1 MSDE-ETVPGTVVGVDSBEGSLHALDWAIDAAAGRGATLRLVYAMGLPLTVTLVGGPIR 59
Qy      49 THPSDDYDRLDLAHERSLREASAVEAAGKLVKLETIDIPRGAPVLYBASRDAEMICV 108
        60 TABSPFVSQAQAKLLLEALRLRVQSAFS---LNAVTEVSAEKAHHLKSAQDAELLV 115
Db      109 GSVGIGRYASSIIGSTATELAERKACPAVAVMRSKVDOPASDINMIIVMTDAPDNEAVLE 168
        116 GSRGYSVGAISLFGSVAQRVASHATCPVVVVPPTSGEAAARGRVVGVDSGEHRAAALR 175
Qy      169 YAAREALRQAPILALGCRPELEIRPDGEF-----RRVQMHHRP 211
        176 FALVEARLRRLAEIVAV-----YAWQAPDAPVDPPTVLQADVAVDREQVAAAREMLLRTV 230
Db      212 DVRYVPTITHTGI-----ARFLADDERVOLAVIGG-GEAGQIARLVGPSGHPVR 261
        231 DEARTPLTORVAVVERPEKPRALALDEGADALIVVSGRGGFTGLLIGSVSQVYN 290
Qy      262 HAECSVLVVR 271
        291 HAVVPVAVVR 300

```

```

RESULT 3
US-11-087-099-6347
; Sequence 6347, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement

```

```

; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 6347
; LENGTH: 152
; TYPE: PRT
; ORGANISM: Streptomyces coelicolor A3(2)
US-11-087-099-6347

```

```

Query Match      10.7%; Score 149; DB 7; Length 152;
Best Local Similarity 29.3%; Pred. No. 6e-06;
Matches 44; Conservative 26; Mismatches 62; Indels 18; Gaps 5;

```

```

Qy      6 EPTMTIIVIGIDGSHAITTAALMGVDEAISRAVPLRLVSYIKPTH-----PS-----PD 54
        5 EPTAR-VVGVDSPPSYALRWNA--DRYARAAG---GVVEAVHWDTFSAGFAPAI 57
Db      55 DYBDLHAERSLREASAVEAAGKLVKLETIDIPRGAPVLYBASRDAEMICVSGVIG 114
        58 DPDFLEQARERPAELEATFPGERPGLKEIIVEGDPSETLIRASQAEIIVGRRGRG 117
Qy      115 RVASSIIGSTATELAERKACPAVAVMRSKVD 144
        118 AFARMLGSVSGRCAGHACPVVVRQETE 147
Db

```

```

RESULT 4
US-11-087-099-11409
; Sequence 11409, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 11409
; LENGTH: 163
; TYPE: PRT
; ORGANISM: Glycine max
US-11-087-099-11409

```

```

Query Match      9.3%; Score 130.5; DB 7; Length 163;
Best Local Similarity 29.3%; Pred. No. 0.00026;
Matches 46; Conservative 25; Mismatches 69; Indels 17; Gaps 6;

```

```

Qy      1 MSGRGEPTMTIIVIGIDGSHAITTAALMGVDEAISRA-VPLRLVSI--KPTSPDDY- 56
        1 MATSGETKQVWVIGIDSDPSFYALQWTLDLHLSPANVPKFKIPLVYAKPSVASAVGFV 60
Db      57 ----DRDLHAERSLRE-AQSAVEAAGKLVK-----IENDIPRGAPVLYEA--SRDA 103
        61 GPAAAEVLPVVEADLTKTAITERATLCKKKSUNDVAVEVLEGDPVNLCEAVEKQA 120
Qy      104 EMICVSGVIGRYASSIIGSTATELAERKACPAVAVMR 140
        121 SMLVGSHGVTGTLKRAVLGSVSDYCAHHAHCTMIVK 157
Db

```

```

RESULT 5
US-11-096-568A-4791
; Sequence 4791, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471

```



```

; SEQ ID NO 4791
; LENGTH: 163
; TYPE: prt
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(163)
; OTHER INFORMATION: Ceres Seq. ID no. 14304185
; US-11-096-568A-4791

```

Query Match	9.3%;	Score 130.5;	DB 7;	Length 163;
Best Local Similarity	29.3%;	Pred. No. 0.00026;		
Matches	46;	Conservative	25;	Mismatches 69;
			Indels	17;
			Gaps	6

Oy 1 MSGREPMKTIIVGIDGSHAATTAALMGDEAISA-VPLRLSVI--PETPSPDDY- 56
::
Db 1 MATSSETKQWVAVIGIDBDFSTYAQLTDLHLSPANPKFKILPLYAKESVASANGFV 60

Oy 57 ---DRDLAAHRSLE-AQSAAVEAGKLK-----IEFDIRGPAPVLVEA---SRDA 103
Db 61 GPGAAEVLPVEADLKTAKITERTETELCKKKSVDVAEEVLEGPPRNVLCAAVEGQA 120

Qy 104 E M I C V G S V G I G R Y A S I L G S T A T E L A E K A H C P V A M R 140
| : | | | : | | | : | | : : :
Db 121 S M L V V G S H G Y G T L K R A V L G S V S D Y C A H H A C T V M I V K 157

```

RESULT 6
US-11-096-568A--4790
; Sequence 4790, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Theory
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 4790
; LENGTH: 165
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(165)
; OTHER INFORMATION: Ceres Seq. ID no. 14304184
US-11-096-568A--4790

```

	Query Match	9.3%	Score 130.5;	DB 7;	Length 165;
	.Best Local Similarity	29.3%;	Pred. No. 0.00027;		
	Matches	46;	Conservative	25;	Mismatches 69; Indels 17; Gaps 6
Oy	1	MSGRGEPPTKTIIVIGIDGSHAAITALMVDKAISRA--PLRLVSVI--KPTRSPDDY-	-56		
Dd	3	MATTSSETKQWNVITGIDDSDFSTYALQLDLHLSPANPKRIFLVYAKPSVASAGVF	62		
Oy	57	--DRDIAHAERSLRE-AQSAAVEAAGKLKV-----ISTDIIPRGAPPLYVA--SRDA	103		
Dd	63	GPGAELVPVEADLRKTAITERATELTCKKSVDNDAVEVLGGPKNVLCGAVERHQD	122		
Oy	104	EMICVGSGVIGRYASSILGSTATELEAKHCPAVNR	140		
Dd	123	SMLVVGSHGVGTLKQAVLGVSDDYCMAHNAHCVTMIYK	159		

RESULT 7
US-11-087-099-505
; Sequence 505, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement

```

: FILE REFERENCE: 38-21(53450) B EP
: CURRENT APPLICATION NUMBER: US/11/087, 099
: CURRENT FILING DATE: 2005-03-22
: NUMBER OF SEQ ID NOS: 12464
: SEQ ID NO 505
: LENGTH: 171
: TYPE: PRT
: ORGANISM: Zea mays
US-11-087-099-505

```

Query Match 9.2%; Score 128.5; DB 7; Length 171;
 Best Local Similarity 34.0%; Pred. No. 0.00042;
 Matches 51; Conservative 14; Mismatches 68; Indels 17; Gaps 7;

QY 7 PTMKTIIVGDSHAITALTALGVDBAISRAVP---LRLVSY-IKPTHS-----PDDY 57
:::|::|
Db 16 PGKMTVVGVDBSEHSHFYALOMALGHFPFPGQPOGYRLVVTATFTAAISNGLAGPGAAD 75

QY 58 RDLAHERSL-REKQSVAPACKLYKIETDIP---RKPAGPVLEYA--SRDAIMICYGS 110
:::|:
Db 76 -VIPVEADIDKRSALVERKAGCLTQSADVAFEALGEDANVCCEAVERRHGEMLVGS 134

```

QY      111 VGIGRYASSILGSTATLEAKAHCPVAVMR 140
          |::|||:||||:
DB      135 HGXGAIKRAVLGSVSDYCAHNAHCTVMTVK 164

```

```

RESULT 8
US-11-096-568A-15853
; Sequence 15853, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Therapy
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 15853
; LENGTH: 171
; TYPE: PRP
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(171)
; OTHER INFORMATION: Ceres Seq. ID no. 12349153
US-11-096-568A-15853

```

Query Match	9.2%	Score 128.5	DB 7	Length 171
Best Local Similarity	34.0%	Pred. No. 0.00042		
Matches	51	Conservative 14	Mismatches 68	Indels 17
			Gaps	7
Qy	7	PTMKTIIVGIDGSHAITPAALMGVDEAISRVP--LRLSVY-IKPTHS-----PDYD	57	
Db	16	PGKMTWVGVDSDEHHSFYALQMALGHFFPPGQOQRYLRVYVYTAKPAAASVGLAGPDAAD	75	
Qy	58	RDLAAHERSL-REASAVEAACKLVKLTIDP---RGPAGVILVEA--SRDAEMICVGS	110	
Db	76	-VLPVYEADIKSALRVEKAGKGLCTQASDAVFEALGEDANVLCSEAVEHRHGAEMLVGS	134	
Qy	111	VGIGRYASSLIGSTATTELEAKHCPYAVNR	140	
Db	135	HGYGAIKRAVLGVSDYCDHHAKCTVMVVK	164	

```

RESULT 9
US-11-096-568A-15854
; Sequence 15854, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides

```

```

; TITLE OF INVENTION: Theby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 15854
; LENGTH: 153
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(153)
; OTHER INFORMATION: Ceres Seq. ID no. 12349154
US-11-096-568A-15854

Query Match          9.0%; Score 125.5; DB 7; Length 153;
Best Local Similarity 34.2%; Pred. No. 0.00066;
Matches 50; Conservative 14; Mismatches 65; Indels 17; Gaps 7;

Qy 11 TIVIGIDGSHAAITPAALMGVDEAISRVP--LRIVSV-IKPTSPD-----PDDYDRDLA 61
Db 2 TMVVGVESEHSFPYALQWALOHFPFPGQPOQYRLVVTAKFTPAASAVGLAIPGAD-VLP 60

Qy 62 HAERSL-REAGSAVEAAGKLVKIETDIP---RGPAGVLYEA---SRDAEMICVSGVIG 114
Db 61 YVEADLRKSALRVVEKAGLCTQASDAVFALBEGDANNVLCVAVRRGAEMLVVGSHGYG 120

Qy 115 RYASSIIGSTATELAERKACPVAVMR 140
Db 121 AIKRAVLGSVSDYCAHNAHCTVMTVK 146

RESULT 10
US-11-096-568A-27471
; Sequence 27471, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 27471
; LENGTH: 174
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(174)
; OTHER INFORMATION: Ceres Seq. ID no. 1817050
US-11-096-568A-27471

Query Match          8.8%; Score 122.5; DB 7; Length 174;
Best Local Similarity 22.5%; Pred. No. 0.0014;
Matches 41; Conservative 25; Mismatches 61; Indels 55; Gaps 5;

Qy 8 TMKTIIVIGIDGSHAAITPAALMGVDEAISRVPRLVS-----VIKPTSPD----- 54
Db 6 SINCVVAVAVDGSSEVMALRMALDN-----LKLSSSSSSSFVLVHQPSPSVAGVS 58

Qy 55 -----DYDRDLAHAERSLREAGSAVEAAGKLVKIETDI 87
Db 59 PGTIFGPGSGLEVPAPFTALIEQHQRITPTILIEHA-----SGICAEKSVSRVNVKTQV 112

Qy 88 PRGPAGVLYEASRD--AEIMCVSGVIGRYASSIIGSTATELAERKACPVAVMRKSVQ 145
Db 113 VIDPPTYKICEAVENLHADLVWGSRAYGRIKRMFLGSVSNYCTNHAHCPVVIIPKEDS 172

Qy 146 PA 147
Db 173 SA 174
```

```

RESULT 11
US-11-096-568A-32572
; Sequence 32572, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 32572
; LENGTH: 171
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(171)
; OTHER INFORMATION: Ceres Seq. ID no. 13593512
US-11-096-568A-32572

Query Match          8.7%; Score 122; DB 7; Length 171;
Best Local Similarity 22.5%; Pred. No. 0.0015;
Matches 41; Conservative 24; Mismatches 59; Indels 58; Gaps 5;

Qy 8 TMKTIIVIGIDGSHAAITPAALMGVDEAISRVPRLVS-----VIKPTSPD----- 54
Db 6 SINCVVAVAVDGSSEVMALRMALDN-----LKLSSSSSSSFVLVHQPSPSVAGVS 58

Qy 55 -----DYDRDLAHAERSLREAGSAVEAAGKLVKIETDI 87
Db 59 PGTIFGPGSGLEVPAPFTALIEQHQRITPTILIEHASQCAE-----KSNVKTQV 109

Qy 88 PRGPAGVLYEASRD--AEIMCVSGVIGRYASSIIGSTATELAERKACPVAVMRKSVQ 145
Db 110 VIDPPTYKICEAVENLHADLVWGSRAYGRIKRMFLGSVSNYCTNHAHCPVVIIPKEDS 169

Qy 146 PA 147
Db 170 SA 171

RESULT 12
US-11-096-568A-27882
; Sequence 27882, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 27882
; LENGTH: 179
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(179)
; OTHER INFORMATION: Ceres Seq. ID no. 2123864
US-11-096-568A-27882

Query Match          8.7%; Score 122; DB 7; Length 179;
Best Local Similarity 23.2%; Pred. No. 0.0016;
Matches 43; Conservative 25; Mismatches 61; Indels 56; Gaps 6;

Qy 8 TMKTIIVIGIDGSHAAITPAALMGVDEAISRVPRLVS-----VIKPTSPD----- 54
Db 173 SA 174
```

```

Db      6 SLNCVVAVDGSEVSEMEALRWALDN-----LKLSSSSSDSSFVLVHPSPSVAAGVS 58
Qy      55 -----DYDRDLAH-----AERSLREAQSAVEAAGLVKIE 84
Db      59 PGTFPGGGLVEPAFTAIEHQKRTITTLIEHSAQIAEKSVSRLVSLSP-----VWVK 114
Qy      85 TDIPRGPAPVLVEASRD--AEMI CVSGSVGIGRYASSILGSTATELAEKACPVAVMRSK 142
Db      115 TQVAVGPDKYKICAEVENTLHADLLVWGSRAVGRIRKMFLLGSVSNVCTHHAHCPVYIIRKX 174
Qy      143 VDQPA 147
Db      175 EDSSA 179

```

```

RESULT 13
US-11-096-568A-15855
; Sequence 15855, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 15855
; LENGTH: 151
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(151)
; OTHER INFORMATION: Ceres Seq. ID no. 12349155
US-11-096-568A-15855

```

```

Query Match      8.6%; Score 120.5; DB 7; Length 151;
Best Local Similarity 33.8%; Pred. No. 0.0017;
Matches 49; Conservative 14; Mismatches 65; Indels 17; Gaps 7;

Qy      12 IIVGIDGSHAATTAALMGVDEAISRVP--LRVSV--IKPTHPS-----PDYDRDLAH 62
Db      1 MIVGVDESEHSFYALQWALQHFPFGGQGYRLVVVTAKPTAASAVGLAGGAD-VLPY 59
Qy      63 AERSL-REAQSAVEAAGLVKIEITDIP-----RGPAPVLVEA--SRDAEM CVSGSVGIGR 115
Db      60 VEA DLKRSALRVEKAKGLCTQASDAVEALEGDAARNVLCEAVERHGAEMLVGSHGYGA 119
Qy      116 YASIIIGSTATLEAKHCPVAVMR 140
Db      120 IKRAVLGSVSDYCAHHAHCTWIVK 144

```

```

RESULT 14
US-11-087-099-11256
; Sequence 11256, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 11256
; LENGTH: 172
; TYPE: PRT
; ORGANISM: Trifolium aestivum
US-11-087-099-11256

```

```

Query Match      8.6%; Score 120; DB 7; Length 172;
Best Local Similarity 31.0%; Pred. No. 0.0023;

```

```

Matches 48; Conservative 19; Mismatches 66; Indels 22; Gaps 7;

Qy      3 GRGEPMTKTIIVGIDGSHAATTAALMGVDEAIS--RAVPLRVSV--IKPT-----HP 51
Db      17 GEGKMTT---VGVDESEHSFYALQWTLHFPFGGQQOQYRLVVVTAKPTAASAVGLAGR 73
Qy      52 SPDD---YDRDLAHAERSLREAQSAVEAAGLVKIEITDIPRGPAPVLVEA--SRDAEM 105
Db      74 GAADVLPFVEADLKR--SLRVIDKAKELCAQVSDAVEFEVGEDARNVLC EAVERHHAEM 131
Qy      106 ICVSGSVGIGRYASSILGSTATELAEKHCPVAVMR 140
Db      132 LVGNHGYGAIKRAVLGSVSDYCTHHAHCTWIVK 166

```

```

RESULT 15
US-11-087-099-7223
; Sequence 7223, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 7223
; LENGTH: 175
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(175)
; OTHER INFORMATION: unsure at all Xaa locations
US-11-087-099-7223

```

```

Query Match      8.5%; Score 118.5; DB 7; Length 175;
Best Local Similarity 26.0%; Pred. No. 0.0032;
Matches 45; Conservative 15; Mismatches 54; Indels 59; Gaps 5;

Qy      7 PYMKTIIVGIDGSHAATTAALMGVDEAI---SRAVPLRVSVIKPTHSPDDYDRDLAH 62
Db      16 PGKRTVTVGVDESEHSFYALQWALQHFPFGAPPAAVPPSSCVTAKPX----- 62
Qy      63 AERSLREAQSAVEAAGK-----LVKIEITDIP----- 89
Db      63 -----AASAVGLAGPYAADVLPVEADLKRSAALRVEKAKGLCTQVLASDAVEALE 115
Qy      90 GPAPVLVEA--SRDAEMICVSGSVGIGRYASSILGSTATELAEKHCPVAVMR 140
Db      116 GDARNVLC EAVERHGAEMLVGSHGYGAIRKRAVLGSVSDYCAHHAHCTWIVK 168

```

```

Search completed: March 23, 2006, 06:29:37
Job time : 6.52731 secs

```

This page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: March 23, 2006, 04:56:29 ; Search time 104.486 Seconds
(without alignments)
1652.628 Million cell updates/sec

Title: US-10-617-038-28

Perfect score: 1990
Sequence: 1 MRDIPILGRAGFVVNVHMS.....QPEPFTTSPQDADRPSDAG 393

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_21:.*
1: geneseqp1980s.*
2: geneseqp1980s.*
3: geneseqp2000s.*
4: geneseqp2001s.*
5: geneseqp2002s.*
6: geneseqp2003as.*
7: geneseqp2003bs.*
8: geneseqp2004s.*
9: geneseqp2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	length	ID	Description
1	1990	100.0	393	8 ADI37307	Adi37307 M. tuberc
2	425.5	21.4	379	5 ABG77902	Abg77902 High leve
3	244	12.3	290	9 ABM96841	Abm96841 M. xanthu
4	174	8.7	346	9 ABM94210	Abm94210 M. xanthu
5	154.5	7.8	438	8 ADN48111	Adn48111 Thermococ
6	143	7.2	409	4 AAB96393	Aab96393 Putative
7	142.5	7.2	177	4 ADM26447	Adm26447 Hyperther
8	140	7.0	591	7 ADB80145	Adb80145 Mycobacte
9	140	7.0	649	5 ABU05373	Abu05373 M. tuberc
10	139.5	7.0	367	7 ABO80543	Ab080543 Pseudomon
11	138.5	7.0	306	6 ABU22341	Abu22341 Protein e
12	136.5	6.9	372	7 ADM26375	Adm26375 Hyperther
13	133	6.7	944	6 ABR55194	Ab55194 Amino aci
14	132	6.6	3745	6 AAR37001	Aae37001 Micromono
15	130.5	6.6	3362	8 ADK16028	Adk16028 Streptomy
16	127.5	6.4	400	6 ABU20199	Abu20199 Protein e
17	126.5	6.4	851	3 AAB29627	Aab29627 Cat flea
18	126.5	6.4	1445	6 AAE32018	Aae32018 Human bio
19	126.5	6.4	3808	8 ADK16029	Adk16029 Streptomy
20	125	6.3	404	4 AAG81206	Aag81206 Mycobacte
21	124.5	6.3	378	4 AAB96754	Aab96754 Putative
22	124	6.2	719	7 ABE87963	Ab87963 Rice abio
23	123	6.2	470	4 AAU48172	Aau48172 Propionib
24	123	6.2	470	6 ABM44691	Abm44691 Propionib

25	122	6.1	1029	8 ADI5058	Adi5058 Pseudomon
26	122	6.1	1033	7 ABO71865	Ab071865 Pseudomon
27	120.5	6.1	457	4 AAG90385	Aag90385 C glutami
28	120.5	6.1	472	9 ABM96727	Abm96727 M. xanthu
29	120.5	6.1	691	7 ABO70760	Ab070760 Pseudomon
30	120	6.0	508	7 ABO61024	Ab061024 Klebsiell
31	120	6.0	574	9 ABM91171	Abm91171 M. xanthu
32	120	6.0	886	9 AAM92782	Aam92782 M. xanthu
33	119.5	6.0	2969	2 AAM56446	Aam56446 Fragment
34	119	6.0	332	9 AAM96092	Am96092 M. xanthu
35	119	6.0	413	6 ABU34130	Abu34130 Protein e
36	119	6.0	811	4 ABG29434	Abg29434 Novel hum
37	117.5	5.9	4635	8 ADX56098	Adx56098 Streptomy
38	117	5.9	707	8 ADK08714	Adk08714 Human pro
39	116.5	5.9	260	6 ABU33936	Abu33936 Protein e
40	116.5	5.9	517	2 AAM34204	Aam34204 Streptomy
41	116.5	5.9	517	2 AAM55805	Aam55805 Streptomy
42	116	5.8	810	9 AAM93215	Am93215 M. xanthu
43	116	5.8	1402	4 AAG81115	Aag81115 Mycobacte
44	116	5.8	1402	4 AAB66459	Aab66459 Protein e
45	115.5	5.8	386	8 ADN47126	Adn47126 Thermococ

ALIGNMENTS

RESULT 1
ADI37307
ID ADI37307 standard; protein, 393 AA.
XX
XX ADI37307;
XX
AC ADI37307;
XX
XX
DT 22-APR-2004 (first entry)
XX
DE M. tuberculosis low oxygen induced antigen Rv2625c SEQ ID NO:28.
XX
DE mycobacterial infection; vaccine; tuberculosis;
XX
KW Mycobacterium tuberculosis; immunisation; antibacterial; gene therapy;
KW low oxygen induced antigen.
XX
KW Mycobacterium tuberculosis.
XX
OS Mycobacterium tuberculosis.
XX
PN WO2004006952-A2.
XX
PD 22-JAN-2004.
XX
PF 08-JUL-2003; 2003WO-DK000477.
XX
PR 13-JUL-2002; 2002DK-00001098.
XX
XX (STAT-) STRATENS SERUM INST.
XX
XX (STAT-) STRATENS SERUM INST.
XX
XX Andersen P, Rosenkrands I, Stryhn A;
XX WPI; 2004-122778/12.
XX N-PSDB; ADI37352.
XX
XX Use of one or more polypeptides or their fragments, which are expressed
XX during the latent stage of the mycobacterial infection, and/or nucleic
XX acids encoding the polypeptides, for a therapeutic vaccine against
XX tuberculosis.
XX
XX Claim 3; SEQ ID NO 28; 76pp; English.
XX
XX The present invention describes polypeptides or their fragments, which
XX are expressed during the latent stage of a mycobacterial infection,
XX and/or nucleic acids encoding the polypeptides, which are useful for
XX creating a therapeutic vaccine against tuberculosis. Also described: (1)
XX a therapeutic vaccine against tuberculosis comprising one or more
XX polypeptides; (2) a method for treating an animal, including a human
XX being, with tuberculosis caused by virulent mycobacteria, e.g. by
XX Mycobacterium tuberculosis, M. africanum or M. bovis; (3) a method for
XX immunising an animal, including a human being, against tuberculosis

CC caused by virulent mycobacteria; (4) a method of diagnosing tuberculosis
 CC caused by virulent mycobacteria in an animal, including a human being;
 CC (5) a method for diagnosing previous or ongoing infection with a virulent
 CC mycobacterium; and (6) a method of diagnosing Mycobacterium tuberculosis
 CC infection in a subject. The polypeptides have antibacterial activities,
 CC and can be used in vaccines and in gene therapy. The polypeptides are
 CC useful for the manufacture of a therapeutic vaccine for treating an
 CC individual who is infected by a virulent mycobacterium, e.g. M.
 CC tuberculosis, and who is not vaccinated with BCG against tuberculosis.
 CC The present sequence represents a low oxygen induced antigen, which is
 CC used in the exemplification of the present invention.

XX Sequence 393 AA;

Query Match 100.0%; Score 1990; DB 8; Length 393;
 Best Local Similarity 100.0%; Pred. No. 4,56-182;
 Matches 393; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRDAIPGRIAGFVNVNHSVLVILMLFTWSLATMPLGTVGGYPAVVYMLGAGAVML 60
 DB 1 MRDAIPGRIAGFVNVNHSVLVILMLFTWSLATMPLGTVGGYPAVVYMLGAGAVML 60
 QY 61 ASLALHELAAHVAVARAGVSVESTLMLFGVTALGGEAKTPKAFRIAFAPATSLALS 120
 DB 61 ASLALHELAAHVAVARAGVSVESTLMLFGVTALGGEAKTPKAFRIAFAPATSLALS 120
 QY 121 ATFGALITTLGAVTRPAIVISVAMWLATVNLGLFNLLPGAPLDGRLVAYLMRRHGD 180
 DB 121 ATFGALITTLGAVTRPAIVISVAMWLATVNLGLFNLLPGAPLDGRLVAYLMRRHGD 180
 QY 181 SVRAGIGARAGRVVALVIALGLAEFVAGVGVVMTLFGMTFPAAREETRISTQQ 240
 DB 181 SVRAGIGARAGRVVALVIALGLAEFVAGVGVVMTLFGMTFPAAREETRISTQQ 240
 QY 241 LFAGRVADAMTAOCHTAPGWINVEDFIQRYVLGSRHSAVPADDSITGLVALRQLRD 300
 DB 241 LFAGRVADAMTAOCHTAPGWINVEDFIQRYVLGSRHSAVPADDSITGLVALRQLRD 300
 QY 301 VAPSRSTTSVGDIALPLHSVPTARPOEPLTALLERMAPLGPSPALVTEGSAVVGIYTP 360
 DB 301 VAPSRSTTSVGDIALPLHSVPTARPOEPLTALLERMAPLGPSPALVTEGSAVVGIYTP 360
 QY 361 SDVARLIDVYRLAOPEPTFTTSPQDADRPSDAG 393
 DB 361 SDVARLIDVYRLAOPEPTFTTSPQDADRPSDAG 393

RESULT 2
 ID ABG77902 standard; protein; 379 AA.
 AC ABG77902;
 DT 29-AUG-2003 (revised)
 DT 05-NOV-2002 (first entry)
 XX High level promoter polypeptide #6.
 XX High level promoter; Cyanobacterium; UV-B light; green alga; plant;
 XX sunlight bioreactor.
 XX Synecocystis sp. PCC 6803.
 XX WO200261098-A2.
 XX 08-AUG-2002.
 XX 30-JAN-2002; 2002WO-US003926.
 XX 30-JAN-2001; 2001US-0264925P.
 XX (DUPO) DU PONT DE NEMOURS & CO E. I.

PI Huang LL, Larossa RA, McCluskey MP;
 XX WPI; 2002-619256/66.
 DR N-PSDB; ABS63259.
 XX

PT Regulating gene expression of a coding region in Cyanobacterium, useful
 PT in identifying highly expressed or UV responsive genes and their
 PT promoters, comprises culturing the transformed cell in log phase or in
 PT the presence of UV-B light.
 XX
 PS Claim 4; Page 49-50; 86pp; English.

CC The invention relates to a method for regulating gene expression of a
 CC coding region in a Cyanobacterium comprising culturing the transformed
 CC Cyanobacterium in the log phase or in the presence of UV-B light, where
 CC the promoter region is activated and the coding region is expressed. The
 CC method is useful for identification of highly expressed genes or UV
 CC responsive genes, and their corresponding promoters. These genes and
 CC promoters are useful for constructing expression vectors in
 CC Cyanobacteria, green algae or plants and for the production of
 CC biomaterials from sunlight. This sequence represents a high level
 CC promoter polypeptide of the invention. (Updated on 29-AUG-2003 to
 CC standardise OS field)

XX Sequence 379 AA;

Query Match 21.4%; Score 425.5; DB 5; Length 379;
 Best Local Similarity 30.0%; Pred. No. 6,96-32;
 Matches 115; Conservative 70; Mismatches 147; Indels 51; Gaps 11;

QY 1 MRDAIPGRIAG--FVNVNHSVLVILMLFTWSLATMPL--PGTVGGYPAVVYMLGAGG 55
 DB 11 MNNNIRVGSLEGFIPFVNPFW--FLILGLVLTSGODIARPPOLSGTP---MILGLIT 64
 QY 56 AVMLLASLHELAAHVAVARAGVSVESTLMLFGVTALGGEAKTPKAFRIAFAPAT 115
 DB 65 ALPLPASVVAHELASHLVALAQGLEKSTILFLFGSLASLEKESNTWQAFVAVIAGP 124
 QY 116 SLALATFGALATTLGAVTRPAIV--ISVAMWLATVNLGLFNLLPGAPLDGRLVAY 173
 DB 125 SLV-----FLGLTVGGIQLPLPVGQAIIGLGMINLALALFLILGPLDGGNVLSKI 179
 QY 174 LMRHGSVPRAGIGARAGRVVALVIALGLAEFVAGVGVVMTLFGMTFPAAREEBE 233
 DB 180 VMOJTGQNKKILILASHVGQGFGLATLALIGSLIINTLPIGSFWTILIGWFLQNA 239
 QY 234 TRISTQQLFAGRVADAMTAOCHTAPGWINVEDFIQRYVLGER--HSAVPVADRG 292
 DB 240 RNAQVKQMEAFTEADNVIFNSPIIPAGLNTREPRANDYVIGKTPWRRFVYAGDN 299
 QY 293 VALRQLRDVAPSRSTTSVGDIALPLHSVPTAR--POEPLTALLERMAPLGPSPAL 351
 DB 300 LATEDIK-----HVPTSDMPQYTVDSLMQY-----PQGMVTYNAN 334
 QY 352 SAVVGIVTPSDVARLIDVYRLAQ 374
 DB 335 QSLF-----EVAQLDQOKLSE 351

RESULT 3
 ID ABM96841 standard; protein; 290 AA.
 AC ABM96841;
 DT 02-JUN-2005 (first entry)
 XX M. xanthus protein sequence, seq id 16040.
 XX Transgenic plant; DNA replication; gene regulation; gene expression.
 XX Myxococcus xanthus.

PN US6833447-B1.
XX
XX 21-DEC-2004.
XX
XX 10-JUL-2001; 2001US-00902540.
XX
XX 10-JUL-2000; 2000US-0217883P.
XX
XX (MONS) MONSANTO TECHNOLOGY LLC.
XX
XX Goldman BS, Hinkle GJ, Slater SC, Wiegand RC,
XX
XX WPI; 2005-028716/03.
XX
XX New substantially purified *Myxococcus xanthus* nucleic acid molecule
XX encoding a nitrite reductase, useful for determining gene expression,
XX identifying mutations in a gene of interest, and for constructing
XX mutations in a gene of interest.
XX
XX Example 2; SEQ ID NO 16040; 25pp; English.
XX
XX The invention relates to a substantially purified nucleic acid molecule
XX encoding a nitrite reductase of SEQ ID NO. 11926. Further disclosed is a
XX recombinant DNA construct for expression of a nitrite reductase gene in a
XX plant cell, and a plant cell comprising the recombinant DNA construct.
XX The nucleic acid is useful for determining gene expression, identifying
XX mutations in a gene of interest, and for constructing mutations in a gene
XX of interest. Sequences given in records for SEQ IDs 9692-16825 represent
XX a group of 7134 *Myxococcus xanthus* proteins and peptides. Note: The
XX sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from USPRO
XX
XX SQ Sequence 290 AA;

Query Match 12.3%; Score 244; DB 9; Length 290;
Best Local Similarity 30.5%; Pred. No. 1.3e-14;
Matches 91; Conservative 48; Mismatches 117; Indels 42; Gaps 11;
QY 88 LFGGVTALGSEAKTEPKAARFAPATSLATFGALATFAGVTPPAIVISVAMWLA 147
DB 2 MGVGVSELTETAPRRRDEALMAAVGLTSLAVLIGATWLLQEVRSNLOFAP-PTMA 60
QY 148 TVNLLGLFNLPGARLDGRLVRAVYLRHGDVSRAGIGARAGVVALVIALGLAEF 207
DB 61 SLNLFGLVFNLPPAPMDGRTVRASLAGRLG-WVRATQVASMGRGFAVLF----- 111
QY 208 VAGGLVGWV-----LAFIGMTFPAAREEETRISTQULFAGVRVADANTAOPTAP 259
DB 112 -----GVMAVLSLNPFVVIAPFTFNGAGCEAOQVKKTLERVPVADLMTTPRRVGD 164
QY 260 GMINVEDFI-----QRYVLSGERHSAYPVADRDSITGLVALRQLRDVA PRSRSTSVGDI 314
DB 165 AGASIEQALMDLRRLRL-----LPYTE-DERPVGVSLETVAAPDSESRMTTIVEV 217
QY 315 ALPLHSVTPARPOBELTALLERMAPLGRSRPALVTESAVV-GIYTPSDVALIDVVR 371
DB 218 MVP--AVVVRUDEDGWTN-LRRMAR-EERPOLVVEADGVLAGTIDVNDVGRGMAVLYQ 271

RESULT 4
ABM94210
ID ABM94210 standard; protein; 346 AA.
XX
XX ABM94210;
XX
XX 02-JUN-2005 (first entry)
XX
XX M. xanthus protein sequence, seq id 13409.
XX
XX Transgenic plant; DNA replication; gene regulation; gene expression.
XX
XX Myxococcus xanthus.
XX

PN US6833447-B1.
XX
XX 21-DEC-2004.
XX
XX 10-JUL-2001; 2001US-00902540.
XX
XX 10-JUL-2000; 2000US-0217883P.
XX
XX (MONS) MONSANTO TECHNOLOGY LLC.
XX
XX Goldman BS, Hinkle GJ, Slater SC, Wiegand RC,
XX
XX WPI; 2005-028716/03.
XX
XX New substantially purified *Myxococcus xanthus* nucleic acid molecule
XX encoding a nitrite reductase, useful for determining gene expression,
XX identifying mutations in a gene of interest, and for constructing
XX mutations in a gene of interest.
XX
XX Example 2; SEQ ID NO 13409; 25pp; English.
XX
XX The invention relates to a substantially purified nucleic acid molecule
XX encoding a nitrite reductase of SEQ ID NO. 11926. Further disclosed is a
XX recombinant DNA construct for expression of a nitrite reductase gene in a
XX plant cell, and a plant cell comprising the recombinant DNA construct.
XX The nucleic acid is useful for determining gene expression, identifying
XX mutations in a gene of interest, and for constructing mutations in a gene
XX of interest. Sequences given in records for SEQ IDs 9692-16825 represent
XX a group of 7134 *Myxococcus xanthus* proteins and peptides. Note: The
XX sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from USPRO
XX
XX SQ Sequence 346 AA;

Query Match 8.7%; Score 174; DB 9; Length 346;
Best Local Similarity 27.4%; Pred. No. 6.6e-08;
Matches 92; Conservative 37; Mismatches 137; Indels 70; Gaps 17;
QY 58 MLASLALHAEHLAHVVARBAVSVESTLWTFGVTALGSEAK-TPKAARFAPATPS 116
DB 19 IVPSVLIHELGHALAPRRYGCPR-IEIHMGGTTCQHDAAHLTHQOSARVAFGRPGIG 77
QY 117 LALSATFGALA-----ITLAGVTPPAIVISVAMWMTATVNLGLFNLPGAPLDG 166
DB 78 FLAGGLTWGLSQVLPLBPGGLADQVR-----QFLM-----VNIGMGLFNLPMQPLDG 127
QY 167 GR-----LYRAYLMRRHGDVSRAGIGARAGVVALVIALGLAEFVAGLVGVTWLAFIG 222
DB 128 GHLLADLVRAASGYRRHGV-LGVG-----IATAVVVIGLAIW-SKQIMMGMLAMVLG 178
QY 223 WFTPAAREEETRISTQULFAGVRVADANTAOPTAP-GWINVEDFIQRYVLSGERHSAYP 281
DB 179 VMLIEQLRRTTKRPAERFALPR---LVKKPRAAPGALISIQLMDEL---RGTRRP 230
QY 282 --VADRDSITG-----LVALLRQLRDVA PRSRSTSVGDIALPLHSVTPARPOBEL----- 330
DB 231 PGADADDDLEDGPRDPLVGEMLDNGLP-----LAVGSLSAFTQAPLARTGALVIAL 286
QY 331 --TALLERMAPLGRSR-----ALVTESAVV 355
DB 287 LHTGRLOGLASILDSSHARQLSEDTLALISRHACTV 322

RESULT 5
ADN48111
ID ADN48111 standard; protein; 438 AA.
XX
XX ADN48111;
XX
XX 01-JUL-2004 (first entry)
XX
XX Thermococcus kodakaraensis KOD1 protein sequence SeqID1989.
XX
XX

Db 341 TLGFLTLGLAYFWPGMILMGLILMLGRVGNFGALDEVSPILTSR 365

RESULT 7

ADM26447 standard; protein; 177 AA.

ADM26447;

20-MAY-2004 (first entry)

Hyperthermophile Methanopyrus kandleri protein #1053.

hyperthermophile; protein stability enhancement;

protein activity enhancement.

Methanopyrus kandleri.

WO2003076575-A2.

18-SEP-2003.

04-MAR-2003; 2003WO-US006664.

04-MAR-2002; 2002US-0361742P.

14-MAY-2002; 2002US-0380423P.

16-SEP-2002; 2002US-0410974P.

(FIDE-) FIDELITY SYSTEMS INC.

(MALY/) MALYKH A.

Sleashev AI, Pavlov A, Pavlova N, Kozayvkin S;

WPI; 2003-748383/70.

N-PSDB; ADM27081.

New isolated nucleic acids encoding any of about 1700 Methanopyrus

kandleri proteins, and the encoded proteins, useful as a medicaments or

as diagnostic agents.

Claim 31; SEQ ID NO 1053; 1023bp; English.

The invention comprises the amino acid sequence of proteins from the

hyperthermophile Methanopyrus kandleri, the invention also comprises the

complete genome from Methanopyrus kandleri. The Methanopyrus kandleri

proteins of the invention are useful for enhancing the stability and/or

activity of other proteins. The Methanopyrus kandleri genome is useful in

a variety of diagnostic and analytical methods. The present amino acid

sequence represents a Methanopyrus kandleri protein of the invention.

Sequence 177 AA;

Query Match 7.2%; Score 142.5; DB 7; Length 177;

Best Local Similarity 31.3%; Pred. No. 3.7e-05;

Matches 56; Conservative 19; Mismatches 77; Indels 27; Gaps 9;

25 LMLFTWMSLTLMLPGVGGVPAVVWLL-GAGGAVMLLSLAHELHNAVVARAGVSBS 83

19 LGLMWGTT---AKTGFFSLTLHSLIFGPIGVAIVATLF-HELGHYLACKLIDVPASP 73

84 VTLMLFGVTLALGSAKTPKAFRIAPGATSLALSTFGALATTLGAVTPAIVISVA 143

74 PVFPLGLGAFVGHGWTDPDKVFP-IAAGP-----FGG--ALAGIPILVYIPKKA 120

144 WMLATVNLILGLFNLIPGAPLDGRLVRAVLMRRHGDVRAIGTAARGRVVALVIL 202

121 VW---NGIQLFNLIPPLDGSIVLRG-LWM---PSPTAMIGVATVLAALSLVGI 171

RESULT 8

ADB80145 standard; protein; 591 AA.

XX ADB80145;

04-DEC-2003 (first entry)

Mycobacterium tuberculosis nutrient starvation-inducible protein #92.

mycobacteria; nutrient starving condition; mycobacterial latency;

mycobacterial infection; vaccine.

Mycobacterium tuberculosis.

WO2003004520-A2.

16-JAN-2003.

04-JUL-2002; 2002WO-GB003052.

04-JUL-2001; 2001GB-00016385.

05-OCT-2001; 2001GB-00023993.

(MICR-) MICROBIOLOGICAL RES AUTHORITY.

James BM, Marsh P, Hampshire T;

WPI; 2003-210338/20.

N-PSDB; ADB80146.

New mycobacterial peptide, useful for the manufacture of a medicament for

treating or preventing, or a diagnostic reagent for identifying,

mycobacterial infection.

Claim 2; Page 312-315; 442bp; English.

The invention comprises the amino acid and coding sequences of

mycobacterial (Mycobacterium tuberculosis) proteins which are up-

regulated under nutrient starving conditions and maintain mycobacterial

latency. The DNA and protein sequences of the invention are useful for

the treatment, prevention and diagnosis of a mycobacterial infection. The

present amino acid sequence represents a Mycobacterium tuberculosis

protein of the invention.

Sequence 591 AA;

Query Match 7.0%; Score 140; DB 7; Length 591;

Best Local Similarity 24.0%; Pred. No. 0.00032;

Matches 111; Conservative 49; Mismatches 169; Indels 134; Gaps 23;

10 IAGFVNVH---WS---VVIWLF-----TWSLATMLPGTVG---GYP 44

70 VAGAIIRVAPLWMTGPAISLVVLQLLESLLALRALVYISWRPVILIPITPALTPLAVP 129

45 AVVYVMLGAGAVNMLLSLAHELHNAVVARAGVSBSVTMLFGVTLALGSAKTPKA 104

130 GFVWMAALNSLPMLAA--LAMVCADAILLVRTGNHRAVAV---GVLVYIGLLFEKA 183

105 AF--RIAPGATSLALSTFGALA--ITLAGVR--TPAIVISVAM---WLATV----- 149

184 AVIPFVSFAVVALQCHVRGDRSALATVWRAGVRLMTSIALTVGMVALYLAIVDQRRSS 243

150 -----NLLGLFNLIPGAPLDGRLVRAVLMRRHGDVRAIGTAARGRVVA-- 196

244 DLSMTWDLTCSRVTNGIYIPALAGGPMWARAPASP-----ATPPAVVWL 290

197 --LTLALGLAFPAAGLVGVWMLAFIAGVTFEFAAREEFTISTQQLFAGRVADAMTAQ 254

291 GMLVLIANLALSLVRKRIGIPWMLTAAG--YAAACQPIFLMRSSPPTALELQTLAYF 347

255 PHTAPGMINVEDFIQRYVLGERHSAYPVADRDSITGLVALRQLRDVAFRRSTTSVGD 314

348 P-----DLY--VVL-----ALLNAVHQQ-----APNRAGTRWD-- 374

315 ALPHSVFT-ARPEPLTALIERNAPGPRRALVTEG--SAVVGIVTSDVARLID-- 368

```
Db      375 ASPARAATVSAVLAFLTSSLYSTATFLASWDFNETBEYLKXQAASLAAASGAPLLDQE 434
Qy      369 -----VYRLAQPE-----PFTTSPQDADRFSDAG 393
Db      435 VDPVLQGVAMPENLASHMFALLRVPRPFATTTQLRMFSTGTG 477

RESULT 9
ABU05373
ID      ABU05373 standard; protein; 649 AA.
XX
AC      ABU05373;
XX
DT      08-APR-2003 (first entry)
XX
DE      M. tuberculosis and M. leprae marker protein #24.
XX
KM      Mycobacteriosis; survival; virulence; protective antigen; vaccine;
XX      mycobacterial disease; tuberculosis; leprosy.
XX
OS      Mycobacterium tuberculosis.
OS      Mycobacterium leprae.
XX
FN      WC200274903-A2.
XX
PD      26-SEP-2002.
XX
PF      22-FEB-2002; 2002WC-IB001973.
XX
PR      22-FEB-2001; 2001US-0270123P.
XX
PA      (INSP ) INST PASTEUR.
XX
PI      Cole S;
XX
XX      WPI; 2002-755985/82.
XX
PT      Identifying and selecting genes for survival or virulence of mycobacteria
PT      by a comparative genomic analysis of the sequences of Mycobacterium
PT      tuberculosis and M. leprae.
XX
PS      Claim 17; Page 190-192; 874pp; English.
XX
XX      This invention relates to a novel method for identifying essential genes
XX      for survival or virulence of mycobacteria species. The method comprises
XX      aligning the genomic sequence of a first mycobacterium species on a
XX      genomic sequence of a second mycobacterium species and selecting a
XX      polynucleotide sequence that is highly conserved in both genomes with no
XX      counterparts in other bacterial genomic sequences and that corresponds to
XX      an essential gene for the survival or virulence of mycobacterium species.
XX      The method of the invention is useful for detecting M. tuberculosis or M.
XX      leprae infection. The method reduces the number of potential new targets
XX      and protective antigens for new drugs and vaccine compositions to treat
XX      and prevent mycobacterial diseases, particularly tuberculosis and
XX      leprosy. The present sequence represents a marker protein from
XX      Mycobacterium tuberculosis and Mycobacterium leprae identified using the
XX      method of the invention
XX
SQ      Sequence 649 AA;
Query Match      7.0%; Score 140; DB 5; Length 649;
Best Local Similarity 23.5%; Pred. No. 0.00037;
Matches 114; Conservative 62; Mismatches 172; Indels 138; Gaps 25;
```

```
Qy      85 TLMFGSVTALGGRAPKAPKAFRIAFAG---PATSLALSAT---FGALATTLGAVRTPAI 138
Db      225 TCWLRALIVVAGGSLPVAYOLFRMGYGLLVPGTALAKDAADKMSGGITLYLSNFNPQYV 284
Qy      139 VISVAMWLTATVNL-LGLFNL-----PGAPLDGCR-----L 169
Db      285 L-----WPLVLVLGLLMLLTHRWPSFMHPLETPESSGVAARAVOSPAAVVFVPSGL 339
Qy      170 VRAYLMRR-----HGDSVRAGIGAAARAGVVALVLIALLGLAEF--VAGLVGGV---W 217
Db      340 LQAFYVIRGGDFMGHGVLLAPFLCLLAPVVIIPVISEG-ADFSRQTGMWLAGVTSLLW 398
Qy      218 LAFIWFIFPA-----AABEETRISTOOLPAGVAVADMTAP--HTPAGININDEFIQRY 271
Db      399 LGVAGMSLMAANSFGWGDATNVS---YSGIDERRFVQATGAHP--LTADYV--- 449
Qy      272 VLGERHSAYVADRDGSIITGLVALRQUDVAPRPSRTSTVGDIAPHSVPTARPOEPLT 331
Db      450 -----GYP---RMAAV--LVALLNTPDQALLPSGNVIKMDLVPMIQSPSSPSPD 497
Qy      332 ALLERMAPLGPGRBALVTESAVVGIVTPSDVARLIDVYRLAQPEPTTSPQ-----D 385
Db      498 SLVSGK----PQHTVFPT-NIGMLGMNVGLDV-RVIDQIGLAPLAQHTERLQHGRIQHD 551
Qy      386 ADRFSD 391
Db      552 KNLFPP 557

RESULT 10
ABO80543
ID      ABO80543 standard; protein; 367 AA.
XX
AC      ABO80543;
XX
DT      29-JUL-2004 (first entry)
XX
DE      Pseudomonas aeruginosa polypeptide #12718.
XX
KM      Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
XX
OS      Pseudomonas aeruginosa.
XX
FN      US6551795-B1.
XX
PD      22-APR-2003.
XX
PF      18-FEB-1999; 99US-00252991.
XX
PR      18-FEB-1998; 98US-0074788P.
XX      27-JUL-1998; 98US-0094190P.
XX      (GENO-) GENOME THERAPEUTICS CORP.
XX
PI      Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX      WPI; 2003-615309/58.
XX
DR      N-PSDB; ABD14114.
XX
PT      Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT      useful as molecular targets for diagnostics, prophylaxis and treatment of
PT      pathological conditions resulting from bacterial infection.
XX
PS      Disclosure; SEQ ID NO 29289; 455pp; English.
XX
XX      The invention relates to Pseudomonas aeruginosa polypeptides and the
XX      polynucleotides encoding them. The sequences are useful in diagnosis and
XX      therapy of pathological conditions, as molecular targets for diagnostics,
XX      prophylaxis and treatment of pathological conditions resulting from a
XX      bacterial infection, for evaluating a compound, such as a polypeptide,
XX      for the ability to bind a P. aeruginosa nucleic acid, as components of
XX      effective antibacterial targets, as targets for antibacterial drugs,
XX      including anti-P. aeruginosa drugs, as templates for recombinant
```

CC production of *P. aeruginosa*-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of *P. aeruginosa*-caused
CC infection, and in detection of *P. aeruginosa* sequences or other sequences
CC of *Pseudomonas* species using biochip technology. Sequences AB067826-
CC AB084396 represent *P. aeruginosa* polypeptides of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html

SQ Sequence 367 AA;

Query Match 7.0%; Score 139.5; DB 7; Length 367;
Best Local Similarity 24.9%; Pred. No. 0.00019;
Matches 98; Conservative 48; Mismatches 143; Indels 105; Gaps 18;

QY 24 ILMLPTWMLATMLPTGVGYPAVVY-----WLLGAGVAMLLASLLAHLAAVVARAG 78
DB 27 VTKRLYV--ATPLVGNALGMLASLVGWPLPMWIGSLVLMVLRCLADPLAEVPAKRCG 84
QY 79 VSVESVTLMFG-GVTALGGEAKTPKAAFRFAAGPATSLASA--TFGALATTLAGVRT 135
DB 85 -----QWVGIGI-----GLHFTPAVIEQV-----LAHSVITIVGAVATTLSSVLA 125
QY 136 PAIVSVAMWMLATVNLGLFNLFGAPLDGSLVRAVLMRRKGSVYAGIGAAAGV 195
DB 126 IAFMRSGEDRATAPF-----ASMPGASBMVNLGQRHG-AVLSRVAAAGSLRL 174
QY 196 ALVITLGLAEFVAGLVGV-----WLA-----IGMFIFAAREBEET-- 235
DB 175 LVTLVLPVPAFQVLGGGQGPQHPAAPVDMWMLALLFPAGALVALGW-----OKLRQ 226
QY 236 ----ISTOOLPAGVVA--DAMTAPHTAG--WINEVDPIQRVYVIGRHSAYPVADRQ 287
DB 227 NPMILGPLLAAVGLGFDLHGLPAGSSGVQWL-----IG 263
QY 288 SITGVALRQLDVAPSRRTTSVGDIALPLHSVPTAPQEPLT-----ALLERNAPLGP 342
DB 264 SALGCHFNRSFPRSPAPAVSRVLVCTLMWMPAALAAHLGLWLTLDHQSILMGMPG31 323
QY 343 RSRALVTGSAN-VGIVTPSDVARLIDVYRLAQP 375
DB 324 ABLSLTAELALQISVPLVTLALQVLRLLVLFLEAP 357

RESULT 11

ABU22341
ID ABU22341 standard; protein; 706 AA.

XX AC ABU22341;

DT 19-JUN-2003 (first entry)

DE Protein encoded by prokaryotic essential gene #7866.

XX KW Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX OS Burkholderia mallei.

XX EN WO200277183-A2.

XX PD 03-OCT-2002.

XX PF 21-MAR-2002; 2002MO-US009107.

XX PR 21-MAR-2001; 2001US-00815242.

XX PR 06-SEP-2001; 2001US-00948993.

XX PR 25-OCT-2001; 2001US-0342923P.

XX PR 08-FEB-2002; 2002US-00072851.

XX PR 06-MAR-2002; 2002US-0362699P.

XX PA (ELIT-) ELITRA PHARM INC.

XX PI Wang L, Zamudio C, Malone C, Hasebeck R, Ohlsen KL, Zyskind JW;

PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX WPI; 2003-029926/02.
XX N-PSDB; ACA26211.

PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.

PS Claim 25; SEQ ID NO 50265; 1766bp; English.

XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences

SQ Sequence 706 AA;

Query Match 7.0%; Score 138.5; DB 6; Length 706;
Best Local Similarity 22.2%; Pred. No. 0.00057;
Matches 110; Conservative 48; Mismatches 159; Indels 179; Gaps 21;

QY 7 LGRIAGFVNVY---HWSVLVTLMLFTWMLPTGVGYPAVVYVLLGAGVAMLLAST 63
DB 121 LGLIAMLVPLAAGHASY-PLTFVCGLAACVSAAGMGRATPLRLILAGSVCMLLFSA 179
QY 64 LABEL-----AAHVARRAGSVESVTLMFGCVTALC-----GE 98
DB 180 VTTLLAFPEQTVVAGSLWAGSLYOPGAAGLRDALMLVPLAALPLVIRPLDPLAGD 239
QY 99 AKTPKAAFRFAAGPATSLA---LSATFGALATTLAGVTPAVYISVAMWMLATVNL-- 152
DB 240 DAAAAGVRV---DARTLAGTVVAVGPAVASIAG--PLSYTG---LIAPNLIRQM 288
QY 153 -----LGLFNLLPGAPLDGRLV----- 170
DB 289 RGAKASRLGA--LVPLAALAGALVVTDSAVLALGIDATLSTGVATAFVGTPLMLAMIR 346
QY 171 PAVLYW---RRHGDVSRAGT-GAAPAGR-----VALVITLALGLAEFVAGLVGVWL 218
DB 347 RGAAMSGAQQPAHVATGTRGATLVRLAVLPMPAVAVLILLAGAAGVILVVGASFGPWI 406
QY 219 AFIGMFIFAAREBEET-----ISTQQLF----- 242
DB 407 GGRWFPAALQRDLARTVLDLRAPRLCALLAGALLGASGVLMQSTIRVPLAGPEVLGV 466

Db 529 RAIPPKGITV-LVGGTALALDLSIHGLFAKMP--LMVVIILTTITVIMFLAFGSVLP 584
 Qy 130 LAGVTPAIVISVAMMLATVNLGLF-----NLIP---GAPLDG----- 166
 Db 585 IKATLMSALTLG-----STMGLITWIFVDGHSKMLNFTPTPLTPKPVGLIILAVFGSLT 639
 Qy 167 -----GRLVAVYIMRRHDSVRAG--IGAARAGRV--ALVILALGLAEFVAGLVGG 215
 Db 640 DVEVFLVSRMVEA---RRGMSSTQGAIRIGTAATGRITTAALIVAVVAGAFVFSDLVMM 696
 Qy 216 VWLAFIGFIFPAAAEETRISTOOLFAGVRYADAMTQPHRAPW-----IN 263
 Db 697 KYLA---GLMAALLDLDAIVV---RMFLVPSVMKLGDDCWMPAPRMAPRLQTRIGLGIH 750
 Qy 264 VEDFIORYVYLGERRHAYPVADRDSITGLVALRQLRDVAPSRSTVSVDIALPLHSVPT 323
 Db 751 LPDEKRRVSNRGRPARPVT-----AGLVAKAAGDRPRPHDPHPLAESRPARSSPA 804
 Qy 324 ARPOEPLTALLERMAPLGRSRALVTESAVVGIPTSDVABLIDVRLAQ-----PEPT 378
 Db 805 SSPB--LTPALEATAPAPASGASTTR--MQIGSSTPPTTRLAAAGRSVGSFASTPPT 860
 Qy 379 FT-----TSPQADRFSDA 392
 Db 861 PTSPAPASGQTRAMPPLAARSTDA 885

RESULT 14

AAE37001 standard; protein; 3745 AA.

AAE37001;

23-OCT-2003 (revised)
07-AUG-2003 (first entry)

Micromonospora carbonacea polyketide synthase (PKS) type I #7.

Macrolide; rosamycin; polyketide; polyketide synthase; PKS; enzyme.

Micromonospora carbonacea.

CA2391131-A1.

19-NOV-2002.

26-JUN-2002; 2002CA-02391131.

26-JUN-2001; 2001US-0307629P.

(ECOP-) ECOP1A BIOSCIENCES INC.

Yang X, Staffa A, Farnet CM;

WPI; 2003-343556/33.

N-PSDB; AAD55817.

Novel isolated polypeptide involved in biosynthesis of macrolides by microorganisms, useful for biosynthesis of macrolides by microorganisms, preferably for biosynthesis of rosamycin.

Claim 13; Page 138-151; 206pp; English.

The invention relates to genes and proteins involved in the biosynthesis of macrolides by microorganisms. In particular it relates to the nucleic acids forming the biosynthetic locus for rosamycin (a 16-member macrolide antibiotic) from Micromonospora carbonacea. The invention is useful for the biosynthesis of macrolides by microorganisms. It allows direct manipulation of macrolides and related chemical structures by chemical engineering of the proteins involved in the biosynthesis of rosamycin. It is useful to catalyse certain biochemical reactions, in vitro or in vivo, to direct or enhance the synthesis or modification of a polyketide, polyketide substrate or its precursor. The present sequence

CC is M. carbonacea polyketide synthase (PKS) type I protein. (Updated on 23
 CC -OCT-2003 to standardise OS field)
 XX
 SQ Sequence 3745 AA;

Query Match 6.6%; Score 132; DB 6; Length 3745;
 Best Local Similarity 25.2%; Pred. No. 0.022;
 Matches 107; Conservative 40; Mismatches 174; Indels 104; Gaps 20;

Qy 31 SLATMLPGTGVGVAVVYVWMLGAGVWMLASLL-AHELAHAVVARRAGVES----- 83
 Db 1063 SLMLAEGADARHPAVP---LGLTASLALIGALADAGTQARLMVTRKAVVSSGEVPDA 1119
 Qy 84 --VTLMFGVYAL-----GGEAKTPKAFRIAPAGPATSIALATFGALATLAG--- 132
 Db 1120 GQAGVWGLGRVAALELPDRWGLVDLPALNGERAFQALADVVGSGNGEDQAVAVASGVYG 1179
 Qy 133 ---VTPPAIVISVAMMLATVNLGLFENLFGAPLDGRLVRAVYIMRRHDSVRAGICAA 189
 Db 1180 RRLVRSRATVTSQDMPARGTILVVG--DTGVAALLAGRLI-----GDGAA 1223
 Qy 190 R---AGRVAVLVALGLAEFVA-----GGLVG-----GWLAFIG 222
 Db 1224 HVTLAEPAASTVGLTGADRVALLDCDPSRDALAGLGAIVPTTVVAPPAVALTLA 1283
 Qy 223 W-----FIFAAREEETRISTOOLFAGVRV-ADAMTQPHRAPGMINVEFIORYLGER 276
 Db 1284 ETPPEDFVAIVAAKTTTAVHLDALAAEAELELDAFVVFSSVSGTW-----GGAG 1332
 Qy 277 HSAYPVADRDSITGLVALRQLRDV-----APSRSTSVG---DIALPLHSVPTAR 325
 Db 1333 HGCV--AAGTARLALVEERRARGLPATIAMTWMADATTAAAGQAPDASAGHE-PDTR 1389
 Qy 326 PQEPLTALLER--MAPGPRSRALVTESANV--VGIPTSDV--ARLIDVRLAQPEPTF 379
 Db 1390 AGPDRBLRRGGILPLDPGALDVLKGAVARGEGLTVADVDMARKFVASTTAAKPTTLF 1449
 Qy 380 TTSPQ 384
 Db 1450 DELPB 1454

RESULT 15

ADK16028 standard; protein; 3362 AA.

ADK16028;

03-JUN-2004 (first entry)

Streptomyces halstedii vncinistatin VlnP3 product seq id 6.

cytostatic; gene therapy; gene cluster; antibiotic; vncinistatin;

dTDP-vicenisamine; cancer; VlnP3; polyketide synthase.

Streptomyces halstedii.

US2004053274-A1.

18-MAR-2004.

04-MAR-2003; 2003US-00378083.

04-SEP-2002; 2002JP-00258838.

(TOKD) TOKYO INST TECHNOLOGY.

Kakimura K;

WPI; 2004-247723/23.

N-PSDB; ADK16023.

New gene clusters encoding a vncenisamine glycosyltransferase polypeptide

PT useful in the biosynthesis of vincenistatin using dTDP-vicenisamine as a
PT substrate and for treating cancer.

PS Claim 8; SEQ ID NO 6; 103pp; English.

XX
CC The invention describes a gene cluster comprising a base sequence
CC represented by base numbers 1-64492 of a 64492 base pair sequence (SEQ ID
CC No. 1) encoding proteins required in the production of antibiotic
CC vincenistatin, given in the specification. Also described are the
CC following: a polypeptide comprising a 419 amino acid sequence (SEQ ID No.
CC 4), given in the specification, or an amino acid sequence having at least
CC 90 % homology with SEQ ID NO: 4, and exhibiting enzymatic activity as a
CC glycosyltransferase; a gene encoding a polypeptide of (1); a gene
CC comprising a base sequence represented by base numbers 2790-4049 in SEQ
CC ID No. 1, a base sequence having at least 90 % homology with the base
CC sequence cited above, and encoding a polypeptide exhibiting enzymatic
CC activity as a glycosyltransferase, or a gene comprising a base sequence
CC that hybridises with the base sequence cited above under a stringent
CC condition, and encoding a polypeptide exhibiting enzymatic activity as a
CC glycosyltransferase; a polypeptide comprising an amino acid sequence
CC represented by a 355, 323, 2260, 3362, 3808, 236, 397, 493, 330, 158,
CC 469, 299, 327, 82, 524, 478, 414 or 5836 amino acids (SEQ ID NO: 2-3 or 5
CC -20), respectively, and producing vincenistatin by catalytic reaction of
CC the glycosyltransferase, comprising the polypeptide of (1); using dTDP-
CC vicenisamine as a starting material. The methods and compositions of the
CC present invention are useful for producing vincenistatin in the treatment
CC of cancer. This is the amino acid sequence of a polypeptide synthase
CC encoded by the Vinp3 gene of the Streptomyces halstedii gene cluster
CC comprising genes encoding proteins required in the production of the
CC antibiotic vincenistatin.

XX
SQ Sequence 3362 AA;

Query Match 6.6%; Score 130.5; DB 8; Length 3362;

Best Local Similarity 24.1%; Pred. No. 0.027;

Matches 87; Conservative 52; Mismatches 153; Indels 69; Gaps 14;

QY 47 VYMLGAGGAVVMLASLAEHLAAHVARRAGVS-----VESVTLMLEGGVTA 94
DB 416 VDM---SAGAVELLTE--AREWPTGRRPRAGVSFGSGTNAHYIIEASEFEPSAVER 470
QY 95 LGGEAKTP-----KAAFRIFAPGAPATSLASATFGALATTLGAVRPAVI 140
DB 471 LAGGVTTPPWVLSARSADALRGQERLLSFVSAAGDVSVDAVYSLGVSRAGLEHRGVV 530
QY 141 SVAMWLATVNLGLFNLPGAPLDG-----GRLVRAYLMRRHSDSVRAGIGARAG 192
DB 531 GE---SRAELLALESLSAGVESPGVVTGRVABGR--AFUFTGCG-AQRVGMGRELA- 582
QY 193 RVVALVLTALGLAEFVAGLVGVWLAFIGWFIAPAAREESTRISTOOLFAGVRVADAMT 252
DB 583 --AAPFPAASLBE--TCGLLEBAGVA-VREVLPABEGSAEALLTRTVYAQAALFAVEV 637
QY 253 AQPHTAPGWIVEDFIQRYVUTGERHSAY--PYADRDGSITGLVALRQLDVAPERRSTTS 310
DB 638 ALFRLVESFGVVPDFVAGHSVGEIAAAHVAGFSLIEDVSLVAARGRLMDALPEGGAIVA 697
QY 311 VGDIALPLHSVPTARPOEPLTALLERMAPLGPSSRLVTEGSAVVGIYTPSDVARLIDVY 370
DB 698 V-----QATEBDVALLLEGVE---DASIAIINGPDVAVVSGTEAGVARVVDVL 742
QY 371 R 371
DB 743 R 743

Search completed: March 23, 2006, 05:10:54
Job time : 107.486 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: March 23, 2006, 05:11:18 ; Search time 15.6356 Seconds
(without alignments)
2418.401 Million cell updates/sec

Title: US-10-617-038-28

Perfect score: 1950
Sequence: 1 MRDIPILGRINGFVVNVHMS.....QPEPTFTSPQDADRFS DAG 393

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Fred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1990	100.0	393	2	H70572
2	486.5	24.4	389	2	AE2036
3	477	24.0	399	2	AE2293
4	459	23.1	390	2	D84367
5	425.5	21.4	379	2	S76029
6	418	21.0	362	2	D69291
7	377.5	19.0	341	2	A69209
8	328.5	16.5	339	2	H64348
9	328	16.5	372	2	H95972
10	288	14.5	372	2	A95338
11	173	8.7	365	2	C64421
12	149.5	7.5	284	2	E97054
13	144	7.2	493	2	AD2070
14	143.5	7.2	373	2	D71142
15	143	7.2	409	2	B75010
16	143	7.2	503	2	S77290
17	142.5	7.2	2569	2	T14164
18	140	7.0	591	2	D70562
19	140	7.0	649	2	H86920
20	139.5	7.0	345	2	D83551
21	139	7.0	191	2	C64376
22	136.5	6.9	289	2	G84026
23	134	6.7	2129	2	T14182
24	133.5	6.7	199	2	A69232
25	133	6.7	944	2	C70839
26	132.5	6.7	360	2	C82445
27	130.5	6.6	871	2	AC2417
28	130	6.5	792	2	A84308
29	129.5	6.5	403	2	B83408

30	128.5	6.5	183	2	A72499	hypothetical prote
31	127.5	6.4	288	2	S18438	sporulation protei
32	127.5	6.4	618	2	B87566	voltage gated chlo
33	125.5	6.3	532	2	B87343	ABC transporter, A
34	125.5	6.3	840	2	P84212	hypothetical prote
35	125	6.3	404	2	G70886	probable integral
36	124.5	6.3	378	2	H75213	serine proteinase
37	124.5	6.3	755	2	B75346	probable competent
38	124	6.2	285	2	T34615	NADH2 dehydrogenas
39	123.5	6.2	367	2	B70839	hypothetical prote
40	123	6.2	370	2	H84351	hypothetical prote
41	122.5	6.2	219	1	D70057	conserved hypotet
42	122	6.1	1029	2	D83120	probable RND efflu
43	121.5	6.1	347	2	A95849	probable ABC trans
44	121	6.1	217	1	G70459	conserved hypotet
45	121	6.1	709	2	P75584	hypothetical prote

ALIGNMENTS

RESULT 1									
H70572	hypothetical protein RV2625c - Mycobacterium tuberculosis (strain H37RV)								
C:Species: Mycobacterium tuberculosis									
C:Date: 17-Jul-1998	#sequence	revision 17-Jul-1998	#text	change 09-Jul-2004					
C:Accession: H70572									
R:Coile, S.T.; Broesch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Felwell, T.; Genclis, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.									
Nature 393, 537-544, 1998									
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrett, B.G.									
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome									
A:Reference number: A70500; MUID:98295987; PMID:9634230									
A:Accession: H70572									
A:Status: preliminary; nucleic acid sequence not shown; translation not shown									
A:Molecule type: DNA									
A:Residues: 1-393 <COL>									
A:Cross-references: UNIPROT:O06187; UNIPARC:UPI00000D1035; GB:Z95387; GB:AL123456; NID:9									
A:Experimental source: strain H37RV									
C:Gene: RV2625c									
C:Superfamily: conserved hypothetical protein MU0392; CBS homology									
Query Match	100.0%	Score 1990:	DB 2:	Length 393:					
Query Similarity	100.0%	Pred. No. 2.6e-137:							
Matches 393:	Conservative 0:	Mismatches 0:	Indels 0:	Gaps 0:					
QY	1	MRDIPILGRINGFVVNVHMSVILMLFTWSLATMLPCTVGYPVAVVYMLGAGAVMLL	60						
DB	1	MRDIPILGRINGFVVNVHMSVILMLFTWSLATMLPCTVGYPVAVVYMLGAGAVMLL	60						
QY	61	ASLAEHLAAVVARAGVSVEVTLMFGVTLAGEAKTPKAAFRFAFGPATSLALS	120						
DB	61	ASLAEHLAAVVARAGVSVEVTLMFGVTLAGEAKTPKAAFRFAFGPATSLALS	120						
QY	121	ATFGALATTLTAGVTPPAIVSVAMMLATVNLGLFNLPGAPLDGGTLVAAYMRHGD	180						
DB	121	ATFGALATTLTAGVTPPAIVSVAMMLATVNLGLFNLPGAPLDGGTLVAAYMRHGD	180						
QY	181	SVRAGIGARAGRVAVLALGLAEFVAGGLVGVWMLAFIWFIFAAAREEETRISTQQ	240						
DB	181	SVRAGIGARAGRVAVLALGLAEFVAGGLVGVWMLAFIWFIFAAAREEETRISTQQ	240						
QY	241	LFAGRVADANTAOPTAPGIVNEDFQRYVYLGERSHAYVADRDGSTITGLVALRQIRD	300						
DB	241	LFAGRVADANTAOPTAPGIVNEDFQRYVYLGERSHAYVADRDGSTITGLVALRQIRD	300						
QY	301	VAPSRSTSVGDIALPLSHVPTARPOEPLTALERNAPLGPRAVLTESAVVGIYTP	360						
DB	301	VAPSRSTSVGDIALPLSHVPTARPOEPLTALERNAPLGPRAVLTESAVVGIYTP	360						
QY	361	SDVARLIDVRLAQPETFTTSPQDADRFS DAG 393							

Db 361 SDVARLDIVRLAODEPFTTSSQDADRSPDAG 393

RESULT 2
AE2036
hypothetical protein all1844 [imported] - Noctoc sp. (strain PCC 7120)
C|Species: Noctoc sp. PCC 7120
A|Note: Noctoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C|Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C|Accession: AF2036
R|Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuriiz, T.; Sasamoto, S.; Watanabe, A.; Iriyuchihara, N.; Shimizu, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A|Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena
A|Reference number: AB1807; MUID:21595285; PMID:11759840
C|Accession: AF2036
A|Status: preliminary
A|Molecule type: DNA
A|Residues: 1-389 <KUN>
A|Cross-references: UNIPROT:08VYX5; UNIPARC:UP100000CE21E; GB:BA000019; PIDN:BA073543.1
C|Genetics:
A|Gene: all1844
C|Superfamily: conserved hypothetical protein M0392; CBS homology

Query Match 24.4%; Score 486.5; DB 2; Length 389;
Best Local Similarity 30.7%; Pred. No. 5e-28;
Matches 118; Conservative 78; Mismatches 169; Indels 19; Gaps 7;

QY 1 MRDAIPLRIGAFVNVNVMVSVLVIIMFTWSLATMLPGTVGYP-----AVYTWLGGAG 54
DB 17 MNGTIRAGNLFQIPFYIHPSWFLVGLTWSYS--GGILAFPOLSGRLAV--LGLA 70
QY 55 GAVNNLLASLHLEHAHAAVARAGVSVTLMLPGVTALGGEKTPKAFRIAPAGPA 114
DB 71 TALLPFAVVAHEHCHSPVALRQGINVSTLFTFGGLASLEKESKTPAGAFWALACPL 130
QY 115 TSLASAFGALATLACVGRTPAIVISYVMMWLAATVNLGLFENLLPGAPLDGRLVAYL 174
DB 131 VSLLLCGVVTTTGVTTATVGPLAALIGV---LASVNLALFNLPLGRLDGGNINLAVI 187
QY 175 WRRHDSVRAGIGAAARAGVVALVILALGLAEFVAGLVGVWLAFLIGWTFPAAAREEET 234
DB 188 WKVTNPNPKGVTFPASRVQDIFGFWVAIASGIFPILTFYGSFANVMNLLIGFFLQNAAGNAQ 247
QY 235 RISTQQLAGVAVDAMPAQPHTAGCWINVDFLQRYVL--GERSAYPVADRDGSLTGLV 293
DB 248 FARVQEKLTGLPAADAATVTDSPISAHLSLEFPADQDIIQGNWRRFLVNNAGQLVGAI 307
QY 294 ALQRADVAAPRSRSTTVGDIALPLHSVPTARPOEPLTALLERMAPLGRSRALVTGESA 353
DB 308 ALDDARNRFTISWTTQIQWWRPFGST-TIKSSGPLLEVVQLLEQKLSLPLVILDNGV 366
QY 354 VVGIVTPSDVARLIDVRLAODEP 377
DB 367 LLGILEKAAIITQLQ--NGTQPNP 388

RESULT 3
AE2293
hypothetical protein alix900 [imported] - Noctoc sp. (strain PCC 7120)
C|Species: Noctoc sp. PCC 7120
A|Note: Noctoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C|Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C|Accession: AE2293
R|Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuriiz, T.; Sasamoto, S.; Watanabe, A.; Iriyuchihara, N.; Shimizu, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A|Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena
A|Reference number: AB1807; MUID:21595285; PMID:11759840
C|Accession: AE2293
A|Status: preliminary

A:Molecule type: DNA
A:Residues: 1-399 <KUR>
A:Cross-references: UNIPROT:Q8YD03, UNIPARC:UPI000000CE94F5, GB:BA000019, PTDN:BAB75599.1,
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: alr3900
C:Superfamily: conserved hypothetical protein MJ0392; CBS homology

Query Match 24.0%; Score 477; DB 2; Length 399;
Best Local Similarity 34.1%; Pred. No. 2,5e-27;
Matches 120; Conservative 59; Mismatches 163; Indels 10; Gaps 4;

Oy 25 LMLFTMSLATMTPGV-GGYPAVYVWVLGAGAVMLLASLAEHLAAVVARRAGSVES 83
Db 21 IMFVILGLATINFGVAVGEMGTVTANTAGLIMSLTGLHLSLHBLGSLSARSGIRVNS 80
Oy 84 VTLMTFGGTALGGEAKTEPKAFRIAPAGPATSLASATGALAITTAGVTPAIVISA 143
Db 81 IITLPFGGIALBBSKTCKGAFOVALAGPVSIGL---FLLRISGTVSDSSPVEMMV 137
Oy 144 WMLATVNLGLFNLLPGAPLDGRLVRAVLMRRHDSVRAGIGAAPAGVVALVIALG 203
Db 138 GDLARINTLVALLFNILPGPLPDGGQVLKAALMQITGDPRFOAVHMAAKAGIILGAYALG 197
Oy 204 LA-EFPAAGLVGGWMLAFICGFPTFAAREBEETRISTQQEPCAGRVADAMTAQPHITACWI 262
Db 198 FAVDFETRELVLGLWTALLGWGFGVRNANSYDRVTTLOETLLEKAVADMTRDPFRIDANO 257
Oy 263 INVEDFIORVVLGBRHSAYPVADRDGSIITGLVALROQRDVAPSRSRTSVGDIALPLHSV 322
Db 258 TLRSFADSTILTSTNEVVPFASDGRYRGVMALIEDRLVRSSEMEQTLSIHNPLEIRP 317
Oy 323 TARPEELTALLERMAPLGPBRBALVTEGSAAVVGIVTSPDVARLIDVYRLAQ 374
Db 318 TVAESIAIGVINKLLENQLPRVTVLTPAGAVALGIIRDGI-----VGALAQ 364

RESULT 4
D84367
hypothetical protein Vng2168c [imported] - Halobacterium sp. NRC-1
C:Species: Halobacterium sp. NRC-1
C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #ext_change 09-Jul-2004
C:Accession: D84367
R:NG, W.V.; Kennedy, S.P.; Mahaias, G.G.; Berquiat, B.; Pan, M.; Shukla, H.D.; Laaky, S.
Jung, K.H.; Alam, M.; Freilach, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ehardt, H.; Lowe, T.M.; Little
A:Title: Genome sequence of Halobacterium species NRC-1.
A:Reference number: AB4160; PMID:20504483; PMID:11016950
A:Accession: D84367
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-390 <STO>
A:Cross-references: UNIPROT:Q9HNB6, UNIPARC:UPI00000063A7B; GB:AEO04437; NID:g10581585; P1
C:Superfamily: conserved hypothetical protein MJ0392; CBS homology

Query Match 23.1%; Score 459; DB 2; Length 390;
Best Local Similarity 33.9%; Pred. No. 5e-26;
Matches 132; Conservative 63; Mismatches 152; Indels 42; Gaps 14;

Oy 7 LGRIAGFVNVMHVSVLILMLFTWSLATMLP--GTVGVYP-----AVYYWML 51
Db 6 ICRVVGPIKIDVTFLVLPVFAYILIGIQLEVNVSTLNGAPFAAGLDATALTAGSVRWL 65
Oy 52 GAGAVMLASDLAEHLAAVAVARRAGSVESVTLMPGCVTALGGEAKTPKAAFRIAFA 111
Db 66 GAATAAVGLFVGVLVHELGSHSVAMRYDTIDAITLWLGGVASLTQDPBEWDDELCTALA 125
Oy 112 GPATSLASA-TFGALAITLAGVKTPTAIYSVAMWLATVNLGLFWLLGAPLPDGGRIV 170
Db 112 GPATSLASA-TFGALAITLAGVKTPTAIYSVAMWLATVNLGLFWLLGAPLPDGGRIV 170

Db 126 GPANVALAAVAYGLAVALPSSLDVTRVFG---YLALMNLALAAFNLLPGFPMDCGYL 182

Qy 171 RAYLMRRHSDSVRAGIGARAGRVVALYLIALGLAEFVAGLVGGVWMLAFIGMFIPAAR 230

Db 183 RALLARTR-TPARARFKIAAEVGLKPAFLVGLAGLLSF-----NVILIGVAFYIYGAA 234

Qy 231 EEFRTISTQQLFAGRVADAMT--AQPTAPGMIVNEPFIQRYVLGERHSAYPVADRGS 288

Db 235 GEAGRTANNAAFEGVTADVMTPADVHTVATATASVALDMS-MLEHRTGTYPV-RDAT 292

Qy 289 ITGVLAQLRDVAPRSRSTTSVGD-ALPLHSVP-TARPOEPLALERMAPLGRSRA 346

Db 293 AVGMTLLDARSVRAVERDANKRVADVMSDDVTITPGSDATDALDQEHV-----GRL 347

Qy 347 LVTEGSA-VVGIVTPSDVARRLDVYRLAQ 374

Db 348 LVVDADGEMVGLTRSD---LMDAGGIQ 373

RESULT 5

576029

hypotheetical protein - *Synechocystis* sp. (strain PCC 6803)

C:Species: *Synechocystis* sp.

A:Variety: PCC 6803

C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004

C:Accession: S76029

R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.; O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Matsumoto, A.; Yamada, M.; Yasuda DNA Res. 3, 109-116, 1996

A:Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp.

A:Reference number: S74322; MUID:97061201; PMID:8905231

A:Accession: S76029

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-379 <KAN>

A:Cross-references: UNIPROT:O55518; UNIPARC:UPI0000139445; EMBL:D64006; GB:AB001339; NID A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

C:Superfamily: conserved hypothetical protein M0392; CBS homology

Query Match 21.4%; Score 435.5; DB 2; Length 379;

Best Local Similarity 30.0%; Pred. No. 1.3e-23;

Matches 115; Conservative 70; Mismatches 147; Indels 51; Gaps 11;

Qy 1 MRDAIPGRIAG-FVNVHMSVLIWLFMTSLATML---PGTVGYPAVVMYLLGAG 55

Db 11 MNNNIRVSGSLGIPFYVPSW--FLILGLVLTSLYQDLARFPQLSGGTP---WILGLIT 64

Qy 56 AVMLIASILAEHLAAVVARAGVSSESVTMLFGCVTALGGEAKTPKAFRIAPAGPAT 115

Db 65 ALLLPASVVAHELGHSVALAQGIEVKSITLFLFGGLASLEKESVTPQAFVAVIAGPAV 124

Qy 116 SLAISATFGALAITLAGRTPAIV--ISVAMWLATVNLILGLFNLLPGAPLDGRLVAY 173

Db 125 SLVL-----FLGLITVGTQILPVPAGQAIIGLGMINILALFNILPGLPDGGVNLKSI 179

Qy 174 LMRHGDGSRVAGIGARAGRVVALYLIALGLAEFVAGLVGGVWMLAFIGMFIPAAREE 233

Db 180 VMQITGNONKGIILASRVGGEGWALIGSLGIIINIIPTGSFMTILIGMFLQNGSSA 239

Qy 234 TRISTQQLFAGRVADAMTAPHTAPGMIVNEPFIQRYVLGER-HSAYPVADRGSITGL 292

Db 240 RNAQVKEGMEAFVADAVIPNSPIIPAGLIRREFANDVIGTIPRRRLVIGADQQLGV 299

Qy 293 VALROLRVAPSRKSTTSVGDIALPLHSVPTAR-PQELTALLERMAPLGRSALVTEG 351

Db 300 LATBEEK-----HVPTSDWPQVTVDSIMQY-----PQQWTVYAN 334

Qy 352 SAVVGIVTPSDVARRLDVYRLAQ 374

Db 335 QSLF-----EVAQLLDQKLSE 351

RESULT 6

D69291

conserved hypothetical protein AF0332 - *Archaeoglobus fulgidus*

C:Species: *Archaeoglobus fulgidus*

C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004

C:Accession: D69291

R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson .; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirsch, E.F. ; Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L. Nature 390, 364-370, 1997

A:Authors: Overbeek, T.; Cotton, M.D.; Spriggs, T.; Artlich, P.; Kaine, B.P.; Sykes, S. Smith, H.O.; Weese, C.R.; Venter, J.C.

A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo

A:Reference number: A69250; MUID:98049343; PMID:9389475

A:Accession: D69291

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-362 <KLE>

A:Cross-references: UNIPROT:O29915; UNIPARC:UPI000005715F; GB:AE001082; GB:AE000782; NID A:Superfamily: conserved hypothetical protein M0392; CBS homology

F:299-346/Domain: CBS homology <CBS>

Query Match 21.0%; Score 418; DB 2; Length 362;

Best Local Similarity 30.5%; Pred. No. 4.4e-23;

Matches 113; Conservative 70; Mismatches 162; Indels 26; Gaps 10;

RESULT 7

A69209

conserved hypothetical protein MTH816 - *Methanobacterium thermoautotrophicum* (strain Del

C:Species: *Methanobacterium thermoautotrophicum*

C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004

C:Accession: A69209

R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;

; Qiu, D.; Spadafora, R.; Viscardi, R.; Wang, Y.; Mierbowicki, J.; Gibson, R.; Jivani, N.

ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Neelling, J.; Reeve, J.N.

J. Bacteriol. 179, 7135-7155, 1997

A:Title: Complete genome sequence of *Methanobacterium thermoautotrophicum* Delta H: func

A:Reference number: A69000; MUID:98037514; PMID:9371463

A:Accession: A69209

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-341 <MTH>

A:Cross-references: UNIPROT:O26907; UNIPARC:UPI0000066737; GB:AE000859; GB:AE000666; NID

A:Experimental source: strain Delta H

C:Gene: MTH616
C:Superfamily: conserved hypothetical protein M0392; CBS homology
F:287-334/Domain: CBS homology <CBS>

Query Match 19.0%; Score 377.5; DB 2; Length 341;
Best Local Similarity 30.2%; Pred. No. 3.6e-20;
Matches 111; Conservative 74; Mismatches 147; Indels 35; Gaps 13;

Qy 4 AIPGRIGAFVNVNMSVLVILMLFTWSLATMLPGTVGYPVAVVYMLGAGAVMLASL 63
Db 3 SVKIFSVSGIPIELDFSL-LMLFIYLAVL-----GFLSVNLAVL-----ITLVEPTV 51
Qy 64 LAHELAAVAVARRAGVSVESVTLMFGVTALGGEAKTPKAAFRIFAPGATSLASATF 123
Db 52 VIHELAAHSYVALRFGEVKIKSILLPIGGSVMEBIPRIPOEFMISIGPLTNI-LMALI 110
Qy 124 GALAITLAGVTPAIVISVAMWMLATVNLGLFNLPGAPLDGGVLAVYMRHGDGVR 183
Db 111 TAVPL-LFGKGA--ASFADGFLAVNLMLAIENLIPFMDGGKILPAIILAARL-NYIR 166
Qy 184 AGIGARAGRVAVLVIALGLAEFVAGLVGVGWLAFGWFIFAAAREETRISTOQLFA 243
Db 167 ATRIASNKGKILAVLM-----AVTGLFYNFILLGFIYIGABEGYATLISLLE 218
Qy 244 GVRVADAMTAOPHTAPGMINVEDFIQRYVLGRHSAYVADRDGSTITGLVALRQIDVAP 303
Db 219 GVRVADVMTENPVTLHPMTVKEALD-VMFREKHWGVEPTBA-GEIRGIVTFPHDISDASR 276
Qy 304 SRR-STTSVGDIALPLHSVTPARPOEPLTALLERAPGPRSRALVTGSAVVGIVPSPD 362
Db 277 DLRAEDVMTGDDV-----TVRDEEVTGALKENRNL-QLGRLPVMDGKLTGITSKTD 328
Qy 363 VARLIDV 369
Db 329 IVRTINL 335

RESULT 8
H64348
hypothetical protein homolog M0392 - Methanococcus jannaschii
C:Species: Methanococcus jannaschii
C>Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 09-Jul-2004
C:Accession: H64348
R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake, R.; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weissbrock, K.G.; Merrick, J.M.; Glodek, A.; rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
Science 273, 1058-1073, 1996
A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Frazer, C.M.; Smith, H.O.; Woese, C
A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii
A:Reference number: A64300; MUID:96337999; PMID:8688087
A:Accession: H64348
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-339 <BUL>
A:Cross-references: UNIPROT:Q57837; UNIPARC:UPI000013972E; GB:U67492; GB:L77117; NID:gis
C:Genetics:
A:Map position: FOR355789-356808
C:Superfamily: conserved hypothetical protein M0392; CBS homology

Query Match 16.5%; Score 328.5; DB 2; Length 339;
Best Local Similarity 24.7%; Pred. No. 1.3e-16;
Matches 95; Conservative 86; Mismatches 144; Indels 59; Gaps 14;

Qy 1 MRDAIPGRIGAFVNVNMSVLVILMLFTWSLATMLPGTVGYPVAVVYMLGAGAVML 60
Db 1 MNYSIRLFKINGIPIELH--ITFIFLVVITIGLSINNSI--FWAVL-----ILLF 48
Qy 61 ASLAAHELAAVAVARRAGVSVESVTLMFGVTALGGEAKTPK-AAFRIFAPGATSLAL 119
Db 49 VSVVLAHELGHSHYAKKGVKIKELILPLIGVAMMD--KAPKGEALRIGAGPLVPSII 105
Qy 120 SATGALA-----ITLAGVTPAIVISVAMWMLATVNLGLFNLPGAPLDGGRLVRAVLM 175

Db 106 CIVLLIVSGFDINDG-----YPLLITLSLNLMLGPNLIPAPMOGRILRAILS 158
Qy 176 RRRGDSVRAGIGAAARAGVVALVIALGLAEFVAGLVGVGWLAFICWFIPAAREETR 235
Db 159 KKVGYLSTMT--ANIGKSLALIMLFLGL-----LSNNIILILVLFVYFGAEGESRV 210
Qy 236 ISTOQLAGRVADAMTAOPHTAPGMINVEDFIQRYVLGRHSAYVADRDGSTITGLVAL 295
Db 211 VEVETIFKNIKAKDIMTPNPVYVTPDMSIEFLD-FWLKKHYGYVVE-NGKLVCIGI 268
Qy 296 RQLEDAVPSRSTTSVGDIALPLHSVTPARPOEPL-----TALLERAPGPRSRALVTE 350
Db 269 GNI-----HKSGTVDYM-----EKRVVSEDTDIKEILRKANDRVAVE 311
Qy 351 GSAVGVTPSPDVARLIDVYRLAQ 374
Db 312 GKKLKGIIISKTDIIRAMSIIEAKE 335

RESULT 9
H95972
conserved hypothetical membrane protein SMD20925 [imported] - Sinorhizobium meliloti (str:
C:Species: Sinorhizobium meliloti
C>Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
C:Accession: H95972
R:Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernanc
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A:Title: The complete genome of the 1.683-kb pSymB megaplasmid from the N2-fixing endos
A:Reference number: A95842; MUID:21396508; PMID:11481431
A:Accession: H95972
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-372 <KUR>
A:Cross-references: UNIPROT:Q92UP7; UNIPARC:UPI00000CB741; GB:AL591985; PIDN:CAC49448.1;
R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
peia, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kise, E.; Komp, C.; Lelaure,
hehalt, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A:Reference number: A96039; MUID:21368234; PMID:11474104
A:Contents: annotation
C:Genetics:
A:Gene: SMD20925
A:Genome: Plasmid
C:Superfamily: conserved hypothetical protein M0392; CBS homology

Query Match 16.5%; Score 328; DB 2; Length 372;
Best Local Similarity 29.4%; Pred. No. 1.6e-16;
Matches 113; Conservative 66; Mismatches 156; Indels 50; Gaps 17;

Qy 7 LGRIAGVNVNMS-VVILMLFTWSLATMLPGTVGYPVAVVYMLGAGAVMLASL 65
Db 7 IGTLAGAIRVHVTFALLIMI--W----FMHYRIGGTPA--MEGIATVAVFCVVL 57
Qy 66 HELAAHAAVARRAGVSVESVTLMFGVTALGGEAKTPKAAFRIFAPGATSLASATFCA 125
Db 58 HEFHIAAARRFGIKTDTITLPIGSAVARKERPBEEGQGFVIAIAGPLVNVAAVLA 117
Qy 126 LAIT-----LAGVTPAIVISVAMWMLATVNLGLFNLPGAPLDGGRLVRAVYMRHGD 180
Db 118 ILGSSGMEQIAGVEDPGSFLAR--LAGVNVVFLVIFNMIPAFPMDDGGRVRAALASR--- 172
Qy 181 SVRAGIAAARAGRVVALVIALGLAEFVAG--GLVGSWMLAFICWFIPAAREETRIST 238
Db 173 -----LWRSRATQIAA--TTGQGLA-FVFGVGLFYNNPLIFIAIFYLAATAEAQNAQT 224
Qy 239 QQLPAGRVADAMTAOPHTAPGMINVEDFIQRYVLGRHSAYVADRDGSTITGLVA----- 294
Db 225 REISGSVMISDVMTTEPATIDRSKTIIDAIDT-LIATTOREFVVDAAAGFEGILTRDDM 283

	Beet	Local	Similarity	25.9%	Pred.	No.0.0011:	
	Matches	53;	Conservative	35;	Mismatches	80;	Indels 37; Gaps 7;
Qy	46	VYVYLLDGGAVMLLASL-LHAE	LAAHVAVRARRGVS	SVSLMLFG	GYTALGG-EATRPK	103	
Db	13	ILFPFLGRGIFSVILVITHEFT	HYLTAMFGSGPDI	ILP	GAVALRKLDVAPQK	72	
Qy	104	AAPIAFAGPRTSLASATFGA	LAIITLAVTRPAIVI-----	SVAMMLATV	NLLGLFN	157	
Db	73	EDIITLSLGGPILNILLAVIF-----	YVIFKN	SGRMHYYS	LYLFTNLSLGFN	118	
Qy	158	LLPGAPDGGRLDPAVYLWR	RHGDSVRAGIGARAG	RYVALVLILGLAE	FPAGGAVNG---	214	
Db	119	LIPAPLPDGGRIKMEVLSKK-----	IPRK	NEVSYIYSIVIGIT-F	ICIFSVGLFF	169	
Qy	215	---GVMLAFIGMFTFAAR	AREETRI	236			
Db	170	SVKSIINMLVAVFILFC	SFKERGRM	194			

RESULT 13
AD2070
hypothetical protein alr2114 [imported] - Nostoc sp. (strain PCC 7120)

A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C/Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C/Accession: AD20701
R:Kaneko, T.; Nakamura, Y.; Molk, C.P.; Kuritz, T.; Sasamoto, S.; Matanabe, A.; Iriuchihara, S.;
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A>Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AD2070
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-493 <KUP>
A:Cross-references: UNIPROT:O8YV67; UNIPARC:UPI00000CE30E; GB:BA000019; PIDD:BAB73813.1
A:Experimental source: strain PCC 7120
C/Genetics:
C/Genes: alr2114

[illegible]

RESULT 14
D71142
hypothetical protein PH051 - *Pyrococcus horikoshii*
C/Species: *Pyrococcus horikoshii*
C/Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 09-Jul-2004
C/Accession: D71142
R/Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Halkawa, Y.; Hino, Y.; Yamamoto, S.; Seki,
M.; Ohtuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguch
DNA Ref. 5, 55-76, 1998
A/Title: Complete sequence and gene organization of the genome of a hyper-thermophilic
/Reference number: A71000; MUID:96344137; PMID:9679194

A:Accession: D11142
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-573 <KAM>
 A:Cross-references: UNIPROT:O58089; UNIPARC:UPI0000062DB7; GB:AF000002; NID:g3235129; P:
 A:Experimental source: strain OT3
 A:Note: this accession replaces an interim accession for a sequence replaced by Genbank
 C:Genetics:
 A:Gene: PH0351

Query Match 7.2%; Score 143.5; DB 2; Length 373;
Best Local Similarity 22.0%; Pred. No. 0.004;
Matches 66; Conservative 37; Mismatches 92; Gaps 9;

```

QY 6 PLGRDAGVNVNHSVSLVILMLFTWLSATM-----LPSTGGYPRVAVVYLLGAGAVWLLA 61
   || || || || || || || || || || || || || || || || || || || || ||
Db 89 PLITIIILFVLITBSTFPAGYILSSLYVATLNEMLPGIKVYLLNALAFSLG-----TIS 142
   || || || || || || || || || || || || || || || || || || || || ||
QY 62 SLALHELAAHVVARRAVSVE-----SVTLMLFGGVTAAGGEAKTPKAARIAFGPA 114
   || || || || || || || || || || || || || || || || || || || || ||
Db 143 ILGTHENGHKIAATLHNVKSTFPYFIDPFSGFIGLGAIVIRVKSPIPRNAAVDLGASGPI 202
   || || || || || || || || || || || || || || || || || || || || ||
QY 115 TSLALSTFGALATTLGAVRTPAY----- 139
   || || || || || || || || || || || || || || || || || || || || ||
Db 203 AGLIV-----AIPRTIIIGLKTSIVPVDYLKQGETIYFGTSILFYALTKFVLGNLPQSG 257
   || || || || || || || || || || || || || || || || || || || || ||
QY 140 -----ISVAMWLAIVNLLGLFNLLPGAPLDGGLVAYVLMRRHDSVRAIGIGARAGRV 194
   || || || || || || || || || || || || || || || || || || || || ||
Db 258 IILHPLAVAGC---VGLIVTFPLNIDIPAAQIDGGHIAALM-----PERAHR I 301
   || || || || || || || || || || || || || || || || || || || || ||
QY 195 V--ALVLIAGLAEFVAVGVLGVGWLAFIG-----WTFPA 227
   || || || || || || || || || || || || || || || || || || || || ||
Db 302 LTVYALGRTITGLSTFWFGMWLLMGILLILMRIGNPGALDEVTPTPGRKALAILIWIIFA 361
   || || || || || || || || || || || || || || || || || || || || ||

```

RESULT 15
B75010
hypothetical protein PAB1063 - Pyrococcus abyssi (strain Orsay)
C|Species: Pyrococcus abyssi
C|Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C|Accession: B75010
R|anonymous, Genoscope
submitted to the EMBL Data Library, July 1999
A|Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome structure
A|Reference number: A75001
A|Accession: B75010
A|Status: preliminary
A|Molecule type: DNA
A|Residues: 1-409 <KAM>
A|Cross-references: UNIPROT:Q9YI92, UNIPARC:UP100000034560, GB:AA248288, GB:AL096836, NIDB
A|Experimental source: strain Orsay
C|Genetics:
A|Gene: PAB1063

Query March	7.2%;	Score 143;	DB 2;	Length 409;
Best Local Similarity	20.7%;	Pred. No. 0.0049;		
Matches 59;	Conservative 47;	Mismatches 83;	Indels 96;	Gaps 10;

Qy	11	AGFVVNVHWSVLVLMFTWSLATM-LPGTVGGYPAVVVWLLGAGAVMLLASILAHELA	69
	:	:	
Db	142	AGYILST-----LYVTTLTEELNLPGIKNTYINMLAFSLG-----IISITGTHBMG	186
	:	:	
Qy	70	HAVVARRAGVSVESYT-----LMLFGVYTLGGEATPKAPRIAPGAPRSL	117
	:	:	
Db	187	HKINA-----SINHVKSFPFYFIFPSPFIGTLGAVIRVKSPITPRNHEVDLVGSGPIAGL	241
	:	:	
Qy	118	ALSAFEGALATILAGVTPAIV-----	139
	:	:	
Db	242	LV-----AIPVTIIIGLMSAVVPINYLEKKEITIFGSSLLFYGLMKVLGDLPQNVGII	296
	:	:	
Qy	140	--ISVAVMLATVNLGLFENLLPGARLDGRLYRAYVIMRHGDSVRAIGCARAGRVAL	197
	:	:	
Db	297	HPLAVVAGM-----VGLLVTFNLIPRAAQDDGGHVARALL-----PEKAKHRLVLY	340
	:	:	

Thu Mar 23 10:41:48 2006

us-10-617-038-28.rpr

Page 7

```

Oy 198 VL--IALGLAEVAVAGGLVGWMLAFIGFIPAAAREEETRISTQQ 240
      | : : | | | : : : : : : : : : : : : : : : :
Db 341 TLGFLTIGLAVFWPGWILMGILLMLGRVGNPGALDEVSPITTSR 385

```

Search completed: March 23, 2006, 05:26:35
Job time : 16.6356 secs

This page Blank (uspto)


```

Db 181 SVRAGIGAAAGRVVVALVLIAGLAEFVAGGLVGWVLAFIGWTFPAAREETRISTQQ 240
Qy 241 LFAGRRVADAMTAQPHTPAFCGMINVEDFIQRYVLGERHSAYPVADRDGSIITGLVALROLRD 300
Db 241 LFAGRRVADAMTAQPHTPAFCGMINVEDFIQRYVLGERHSAYPVADRDGSIITGLVALROLRD 300
Qy 301 VAPSRSTSTSVGDIALPLHSVPTARPOEPLTALLERMAPLGRSRALVTBGSAAVVGIVTP 360
Db 301 VAPSRSTSTSVGDIALPLHSVPTARPOEPLTALLERMAPLGRSRALVTBGSAAVVGIVTP 360
Qy 361 SDVARLIDVYRLAQPEPTFTTSPQADRFSDAG 393
Db 361 SDVARLIDVYRLAQPEPTFTTSPQADRFSDAG 393

RESULT 2
ID 006187 MYCTU PRELIMINARY; PRT; 393 AA.
AC 006187; Q7D6V5;
DT 01-JUL-1997 (TReMBLrel. 04, Created)
DT 01-JUL-1997 (TReMBLrel. 04, Last sequence update)
DT 13-SEP-2005 (TReMBLrel. 31, Last annotation update)
DE PROBABLE CONSERVED TRANSMEMBRANE ALANINE AND LEUCINE RICH PROTEIN
DE (Hypothetical protein).
GN OrderedLocustNames=MT2700, RV2625C;
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium tuberculosis complex.
OX NCBI_TaxID=1773;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=H37Rv;
RX MEDLINE=96295887; PubMed=9634230; DOI=10.1038/31159;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C.M.,
RA Harris D.E., Gordon S.V., Sigmeier K., Gae S., Barry C.E. III,
RA Tekela F., Badcock K., Basham D., Brown D., Chillingworth T.,
RA Connor R., Davies R.M., Devlin K., Feltwell T., Gentles S., Hamlin N.,
RA Holroyd S., Hornby T., Jagsels K., Krogh A., McLean J., Moule S.,
RA Murphy L.D., Oliver S., Osborne J., Quail M.A., Rajandream M.A.,
RA Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrett B.G.;
RA "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544 (1998).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CDC 1551 / Oshkosh;
RX MEDLINE=22206494; PubMed=12218036;
RX DOI=10.1128/JB.184.19.5479-5490.2002;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J.D., DeBoy R.T., Dodson R.J., Gwinn M.L., Haft D.H.,
RA Hickey E.K., Kolonay J.F., Nelson W.C., Umayam L.A., Esmolaeva M.D.,
RA Salzberg S.L., Delcher A., Uitterback T.R., Weidman J.F., Khouri H.M.,
RA Gill J., Minkula A., Bishai W., Jacobs W.R. Jr., Venter J.C.,
RA Fraser C.M.;
RA "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL J. Bacteriol. 184:5479-5490 (2002).
DR EMBL; BX842580; CAB0617.1; -; Genomic_DNA.
DR EMBL; AE000516; AAK47016.1; -; Genomic_DNA.
DR PIR; H70572; H70572.
DR TIGR; MT2700; -;
DR Tuberculin; RV2625C; -;
DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000644; CBS.
DR InterPro; IPR008915; Peptidase_M50.
DR InterPro; IPR006025; Pept_M_zn_BS.
DR Pfam; PF00571; CBS; 1.
DR Pfam; PF02163; Peptidase_M50; 1.
DR SMART; SM00116; CBS; 1.
DR SMART; SM00116; CBS; 1.

```

```

DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 393 AA; 41480 MM; 6790E021215B6F98 CRC64;
Query Match 100.0%; Score 1990; DB 2; Length 393;
Best local similarity 100.0%; Pred. No. 6.5e-130; Indels 0; Gaps 0;
Matches 393; Conservative 0; Mismatches 0;
Qy 1 MRDAIPGRIAGFVNVNWSVLTLMFTWSLATMLPGTVGGYPAVYVWLLGAGAVWLL 60
Db 1 MRDAIPGRIAGFVNVNWSVLTLMFTWSLATMLPGTVGGYPAVYVWLLGAGAVWLL 60
Qy 61 ASLIAHEIAHAHVARRAGVSVESTLMLFGCVTLALGSEAKTPKAFRIAFGPTSLALS 120
Db 61 ASLIAHEIAHAHVARRAGVSVESTLMLFGCVTLALGSEAKTPKAFRIAFGPTSLALS 120
Qy 121 ATFGALAITTAGVTPATIVISVAAWMLATVNLGLFNLPGAPLDGRLVRAVYMRHGD 180
Db 121 ATFGALAITTAGVTPATIVISVAAWMLATVNLGLFNLPGAPLDGRLVRAVYMRHGD 180
Qy 181 SVRAGIGAAAGRVVVALVLIAGLAEFVAGGLVGWVLAFIGWTFPAAREETRISTQQ 240
Db 181 SVRAGIGAAAGRVVVALVLIAGLAEFVAGGLVGWVLAFIGWTFPAAREETRISTQQ 240
Qy 241 LFAGRRVADAMTAQPHTPAFCGMINVEDFIQRYVLGERHSAYPVADRDGSIITGLVALROLRD 300
Db 241 LFAGRRVADAMTAQPHTPAFCGMINVEDFIQRYVLGERHSAYPVADRDGSIITGLVALROLRD 300
Qy 301 VAPSRSTSTSVGDIALPLHSVPTARPOEPLTALLERMAPLGRSRALVTBGSAAVVGIVTP 360
Db 301 VAPSRSTSTSVGDIALPLHSVPTARPOEPLTALLERMAPLGRSRALVTBGSAAVVGIVTP 360
Qy 361 SDVARLIDVYRLAQPEPTFTTSPQADRFSDAG 393
Db 361 SDVARLIDVYRLAQPEPTFTTSPQADRFSDAG 393

RESULT 3
ID 00DJC3 SYNEL PRELIMINARY; PRT; 409 AA.
AC 00DJC3;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE T111304 protein.
GN OrderedLocustNames=c111304;
OS Synechococcus elongatus (Thermosynechococcus elongatus).
OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
OX NCBI_TaxID=32046;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=BP-1;
RX MEDLINE=2225144; PubMed=12240834;
RA Nakamura Y., Kaneko T., Sato S., Ikeuchi M., Katoh H., Sasamoto S.,
RA Watanabe A., Iriyuchi M., Kawashima K., Kimura T., Kashiida Y.,
RA Kiyokawa C., Kohara M., Matsunoto M., Matsuno A., Nakashiki N.,
RA Shimo S., Sugimoto M., Takeuchi C., Yamada M., Tabata S.;
RA "Complete genome structure of the thermophilic cyanobacterium
RT Thermosynechococcus elongatus BP-1.";
RL DNA Res. 9:123-130 (2002).
DR EMBL; BA000039; BAC08856.1; -; Genomic_DNA.
DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000644; CBS.
DR InterPro; IPR008915; Peptidase_M50.
DR InterPro; IPR006025; Pept_M_zn_BS.
DR Pfam; PF00571; CBS; 1.
DR Pfam; PF02163; Peptidase_M50; 1.
DR SMART; SM00116; CBS; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 409 AA; 45373 MM; 7C192803184EDA68 CRC64;

```



```
Db 1 MKSSLKIGSVIGIPKILHITFLILPFAVYIPA-----INFPYGFQEVBPASTKMYLS 54
Qy 53 AGAVMLLASLILAHAAVAVARAGVSVESVTLMLPGCVTALGGEAKTPKAAFPRIAPAG 112
Db 55 SLTILILPASLILHELAHASYLAKOYGVIESITLFLPGVSSMEIRPDPOGEAKMAAPAG 114
Qy 113 PATSLASATFGALATITLAG---VETPAI-----VISVAMVLATVNLILGLFNLPGAPLD 165
Db 115 PFTSLVI-----GSICLLLYRYIILSNPALSENVPYLTITWIGANMLVIGIFNLPAFPAMD 170
Qy 166 GGRILVAVLWRRHGDSPVAGIGAAPAGVVALVIALGLAEVAGVIGVWLAFTIGFI 225
Db 171 GGRVLRSPYARRM-SYVATKTSMAAVGRFPAILMAIFGI-----LIGNMIFPLIALFI 222
Qy 226 FAAREEETRISTQOLFAGVRVADAMTAPHTAPGMINVEDFIQRYVLGERHSAYPADR 285
Db 223 YAGSEERSTQASVSLNLIKVDIMTKEVVSPSPMVEDLILQ-FMEKKHMGYPVVE- 280
Qy 286 DGSITGLVALROLRDVAPSRSTTSVCDI-ALPLHSVPT-ARPOEPLTALLERMAPLGR 343
Db 281 SGNLKGITPTFDIQVPTIDRPVWVSDIMTRDISVSDAQASDVLKLVTSKNI----- 335
Qy 344 SRALVTGSAVVGIVTSPDVARLI 367
Db 336 GRVLVIDNGSLVGLSRTDLVRIL 359

RESULT 6
Q8TU08_METAC PRELIMINARY; PRT; 364 AA.
ID Q8TU08;
AC Q8TU08;
DT 01-JUN-2002 (TReMBLrel. 21, Created)
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Hypothetical protein MA0007.
GN OrderedLocustNames=MA0007.
OS Methanosarcina acetivorans.
OC Archaea; Euryarchaeota; Methanomicrobia; Methanosarcinales;
OC Methanosarcinaceae; Methanosarcina.
OX NCBI_TaxID=2214;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C2A / ATCC 35395 / DSM 2834;
RX MEDLINE=21929760; PubMed=11932238; DOI=10.1101/9r.223902;
RA Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., Macdonald P.,
RA Fitzhugh W., Calvo S., Engels R., Smirnov S., Altnoor D., Brown A.,
RA Linton L., McEwan P., McKernan K., Talamas J., Tirrell A., Ye W.,
RA Zimmer A., Barber R.D., Cann I., Graham D.E., Grahame D.A., Guse A.M.,
RA Hedderich R., Ingram-Smith C., Kuetner H.C., Krzycki J.A.,
RA Leigh J.A., Li W., Liu Y., Mukhopadhyay B., Reeve J.N., Smith K.,
RA Springer T.A., Unayam L.A., White O., White R.H., de Macario E.C.,
RA Ferry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.T.,
RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
RA Metcalf W.W., Birren B.;
RT "The genome of Methanosarcina acetivorans reveals extensive metabolic
RT and physiological diversity.";
RL Genome Res. 12:532-542(2002).
DR EMBL; AE010657; AM03461.1; -; Genomic_DNA.
DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000644; CBS.
DR InterPro; IPR008915; Peptidase_M50.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR Pfam; PF00571; CBS; 1.
DR Pfam; PF02163; Peptidase_M50; 1.
DR SMART; SM00116; CBS; 2.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 364 AA; 39671 MW; F902D7538343134 CRC64;

Query Match 24.4%; Score 485; DB 2; Length 364;
```

```
Best Local Similarity 32.5%; Pred. No. 1,5e-25;
Matches 123; Conservative 76; Mismatches 149; Indels 30; Gaps 12;
Qy 1 MRDAIPGRIRIAGFVNVHWSVTLVLMFTWSLA-TMP-GTVGGYPVAVVWMLGAGAVM 58
Db 1 MKSLKIGSVIGIPKILHITFLILPFAVFAINPYPYGFQEVBPAPKAYALSTLVAIL 60
Qy 59 LLASLILHELAHAAVAVARAGVSVESVTLMLPGCVTALGGEAKTPKAAFPRIAPAGATSLA 118
Db 61 LFASILVHELAHASYLAKRYGVNIESITLFLPGVSSMEIRPDPOGEAKMASAPLSL 120
Qy 119 LSATFGALATITLAG-VETPAVIS-----VAMVLATVNLILGLFNLPGAPLDGRLVR 171
Db 121 I-----GGVCLLIYXYIILPNSALSANPVLITWIGANMLILGIFNLPAFPMDGRLVR 176
Qy 172 AYLRRRGDSVPAGIGAAPAGVVALVIALGLAEVAGVIGVWLAFTIGFIAPAAE 231
Db 177 AFYARRR-STIKATQSAANAKFPAILIALFGI-----LVGNLMFPLIALFTYVQASE 228
Qy 232 EETRISTQOLFAGVRVADAMTAPHTAPGMINVEDFIQRYVLGERHSAYPADRDSITG 291
Db 229 EERSTQASVTLNLIKVDIMTKEVVSPSPMVEDLILQ-FMEKKHMGYPVVD-GGSLKG 286
Qy 292 LVALROLRDVAPSRSTTSVCDI-ALPLHSVPT-ARPOEPLTALLERMAPLGRSRLVLT 349
Db 287 IVFTFDIQVSTLDRPVTFRVSDIMTRDISVSDAQASDALKLVTAHNI-----GRVLVI 341
Qy 350 EGSAVVGIVTSPDVARLI 367
Db 342 DNGELVGLSRTDLVRIL 359

RESULT 7
Q8Y0D3_ANASP PRELIMINARY; PRT; 399 AA.
ID Q8Y0D3;
AC Q8Y0D3;
DT 01-MAR-2002 (TReMBLrel. 20, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Alr3900 protein.
GN OrderedLocustNames=alr3900;
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC MEDLINE=21595285; PubMed=11759840;
RX Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakazaki N., Shimo S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
RT "Complete genome sequence of the filamentous nitrogen-fixing
RT cyanobacterium Anabaena sp. strain PCC 7120.";
RL DNA Res. 8:205-213(2001).
DR EMBL; BA000019; BAB75599.1; -; Genomic_DNA.
DR PIR; AE2293; AE2293.
DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000644; CBS.
DR InterPro; IPR008915; Peptidase_M50.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR Pfam; PF00571; CBS; 1.
DR Pfam; PF02163; Peptidase_M50; 1.
DR SMART; SM00116; CBS; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 399 AA; 43105 MW; EC3C4B458F8F200D CRC64;

Query Match 24.0%; Score 477; DB 2; Length 399;
Best Local Similarity 34.1%; Pred. No. 6.1e-25;
Matches 120; Conservative 59; Mismatches 163; Indels 10; Gaps 4;
```

QY 25 LMLFTWLSATMLPGTV -GGYPVAVVWVLGAGGAVMLLASLAHELAHAVARAGVSSES 83
DB 21 LMFVLLGLATLNFVAVYQEMGTAVTAGLIMSLFFGSLVLLHGLSHLSAARSQGIKXNS 80
QY 84 VTLMLPGVTLALGGEKTPKAAFRAPAPATSLSLSTFGALATTLAAGVPRPAIVISA 143
DB 81 IYLPFGGIALAEEBSKTPGKAFQVAILAPLVSIGL--FLLRLRGSTVSDSPSMVW 137
QY 144 WMLATVNLILGLENLLPGALDGRVLRAYLWRRHSDSVRAGIGAAARVVALVILALG 203
DB 138 GDLAINLVALLFNLLPGPLDGGVVKALMLQITGDRQAYHMAKAGQIIGYALALG 197
QY 204 LA-EFVAGGLGVGVMLAFIIGWFIFAAREEETRISTQQLFAGVRVADAMTAQPHAPGI 262
DB 198 FAVDEFTRELVTGLMTALIGMFGVRNANSYDRVTLQETLLREKAVDAMTRDFRVIDANQ 257
QY 263 NVEDTQRYVGERRSAYVADRDGSIYGLVALROLRDVAPRRSTSVGDIATLPLHSVP 322
DB 258 TLRSFADSYLLATSPEYVFAASDGRYRGMVAIEDRLVERSEMETQTLHSLAHPLELP 317
QY 323 TARPQEPILALERMAPLGRSRAVLTGEGSAVVGIVTPSDVARLIDVYRLAQ 374
DB 318 TYAESTALAGVINKLENEQLPRVYVLTTPAGAVAGIIRDGI-----VGLALQ 364

RESULT 8

OSUYV6 HALMA PRELIMINARY; PRT; 418 AA.

AC OSUYV6;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Hypothetical protein.
GN OrderedLocustNames=trmAC2755;
OS Haloarcula marismortui (Halo bacterium marismortui).
OC Halobacteriales; Halobacteriacea; Halobacteriaceae; Haloarcula.
OX NCBI_TaxID=2238;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 43049;
RX PubMed=15520287; DOI=10.1101/gr.2700304;
RA Baliga N.S., Bonneau R., Facciotti M.T., Pan M., Glusman G.,
RA Deutsch E.W., Shannon P., Chiu Y., Meng R.S., Gan R.R., Hung P.,
RA Date S.V., Marcotte E., Hood L., Ng W.V.;
RT "Genome sequence of Haloarcula marismortui: a halophilic archaeon from
RT the Dead Sea."
RL Genome Ref. 14:2221-2234(2004).
DR EMBL; AY596297; AAV47518.1; -; Genomic DNA.
DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000644; CBS.
DR InterPro; IPR008915; Peptidase_M50.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR Pfam; PF00571; CBS; 1.
DR Pfam; PF02163; Peptidase_M50; 1.
DR SMART; SM00116; CBS; 2.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 418 AA; 45115 MW; 2C6256D1E2C227AD CRC64;

Query Match 23.8%; Score 474.5; DB 2; Length 418;

Best Local Similarity 33.7%; Pred. No. 9.5e-25;

Matches 137; Conservative 65; Mismatches 160; Indels 45; Gaps 16;

QY 7 LGRIAGFVNVAVSVLILMLFTWLSAT-----MLPGTVG-----YPAVYVWVLG 52
DB 23 ICSAGGIPQLDUTFLVLPFWMIIGTQIEQTTELTNGTLNAGLDVAVLTGALVWVLG 82
QY 53 AGGAVMLASLALHLELAHAVARAGVSVESTLMLFGVTLALGGEKTPKAAFRIPAPG 112
DB 83 IGAALGLFTGVVLAHELGHSVAIRYGFPIIDSLITMLFGIADLSEMPEDMKDELVAIAG 142

QY 113 PATSLALSAT -FGALATTLAAGVTPAIYIS--VAMMLATVNLILGLENLLPGALDGR 168
DB 143 PIVSIAIGAVGFVAFQIIPSGAGT--IVESTREFLIGYALNALMAAFNMLPGFPMDCGR 200
QY 169 LVRAVLMRRHSDSVRAGIGAAARVVALVILALGAEFVAGGLGVGVMLAFIIGETIPA 228
DB 201 VLRALLARR -SYAATTTIABVGKIFAVFGLGFGI--FV---LGNFLGLAFPIYIG 253
QY 229 AREEETRISTQQLFAGVRVADAMTAQPH--TAPGMINVEDTQRYVGERRSAYVADRD 286
DB 254 AAGSRQTSMAAPFGVAVADVMTPADVTTVADMSVRELIQT -MFRERHTGYV -KRS 311
QY 287 GSITGLVALROLRDVAPRRSTSVGDIATLPLHSVPYTPARPOEPLTALLERMAPLGRS-- 344
DB 312 GEVGLVTLIEDARAAVQEVERBAVYGVDMVT--EITITISPE---TDVMDALTSLOQNSVG 366
QY 345 RALVT -EGSAVVGIVTPSDVARLIDV-----YRLAQEPFTTSPQ 384
DB 367 RLVTVEDGSEFEGLLTRSDIMTALSLIKSSSDYTAIGSESETETVRP 413

RESULT 9

OSNOR6 SYNPE PRELIMINARY; PRT; 381 AA.

AC OSNOR6;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Hypothetical protein.
GN OrderedLocustNames=BYC1914.d;
OS Synechococcus sp. (strain PCC 6301) (Anacystis nidulans).
OC Bacteria; Cyanobacteriia; Chroococcales; Synechococcus.
OX NCBI_TaxID=269084;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PCC6301;
RA Sugita M.;
RT "Complete genome structure of the unicellular cyanobacterium Anacystis
RT nidulans 6301 (Synechococcus sp. PCC6301).";
RL Submitted (DEC-2004) to the EMBL/Genbank/DBJ databases.
DR EMBL; AP008223; BAD80104.1; -; Genomic DNA.
DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000644; CBS.
DR InterPro; IPR008915; Peptidase_M50.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR Pfam; PF00571; CBS; 1.
DR Pfam; PF02163; Peptidase_M50; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 381 AA; 41197 MW; 7341790A361BCECF CRC64;

Query Match 23.7%; Score 471.5; DB 2; Length 381;

Best Local Similarity 32.0%; Pred. No. 1.4e-24;

Matches 113; Conservative 66; Mismatches 131; Indels 43; Gaps 7;

QY 29 TWSLATMLPGTVGVGPVAVVWVLGAGGAVMLLASLAHELAHAVARAGVSVESTLML 88
DB 20 TWS-----FGLMGALVM-----ALLFASVLAHELGHSILARAGRVSITFL 66
QY 89 FGVYTLALGGEKTPKAAFRAPAPATSLALSTFGALATTLAAGVTPAIVISVAMW--- 145
DB 67 FGVVAALIREERTPGCAFVWALAGPLVSFALAL-----LLISQLMWPAG 111
QY 146 -----LATVNLILGLENLLPGALDGRVLRAYLWRRHSDSVRAGIGAAARVVA 196
DB 112 SPAQVLSLNLGRSLSTIAVFNLLIPGLPLDGGVVKALMLKYGDRYRGMVAIEDRLVER 171
QY 197 LVYLAIGL -AEFVAGGLGVGVMLAFIIGWFIFAAREEETRISTQQLFAGVRVADAMTAQ 255
DB 172 AIAVAIGLFSWFLGSGSGVWMLALIGMFGVRNANSYDRVTLQETLLREKAVDAMTRDFR 231
QY 256 HTAPGMINVEDTQRYVGERRS -AYPVADRDGSIYGLVALROLRDVAPRRSTSVGDI 314

Db 232 RVLEBSLTLRQFAELLITEBOEGFAYFVAS--DGRYRGRI SLATLQTERSSQMDRLTLTDL 290
Qy 315 ALPLHSVTPARQOEPLTALBERMAPLGRSRALVTESGAVVGVTPSDVARLI 367
Db 291 AEPFRRLPALTETANLAQALALQTAOPSYYTVLTPSGAVAGI IUDHAYIQAL 343

RESULT 10
O7NKH6_GLOVI PRELIMINARY; PRT; 377 AA.
ID O7NKH6_GLOVI
AC O7NKH6;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Glr1502 protein.
GN OrderedLocusNames=gllr1502;
OS Gloeobacter violaceus.
OC Bacteria; Cyanobacteria; Gloeobacteria; Gloeobacteriales; Gloeobacter.
OC NCBI_TaxID=33072;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=PC 7421;
RX MEDLINE=22977040; PubMed=14621292;
RA Nakamura Y., Kaneko T., Sato S., Mimuro M., Miyashita H., Tsuchiya T.,
RA Saemoto S., Matenabe A., Kawashima K., Kishida Y., Kiyokawa C.,
RA Kohara M., Maksumoto M., Matsuno A., Nakazaki N., Shimo S.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of Gloeobacter violaceus PC 7421, a
cyanobacterium that lacks thylakoids.";
RL DNA Rep. 10:137-145 (2003).
DR EMBL; BA000045; BAC89443.1; -; Genomic DNA.
DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR008915; Peptidase_M50.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR Pfam; PF00571; CBS; 1.
DR Pfam; PF02163; Peptidase_M50; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
KW Complete proteome.
KW SEQUENCE 377 AA; 40561 MW; F37B3BDBD6BEBF CRC64;

Query Match 23.2%; Score 462.5; DB 2; Length 377;
Best Local Similarity 33.4%; Pred. No. 5,8e-24;
Matches 129; Conservative 63; Mismatches 163; Indels 31; Gaps 13;

Qy 2 RDALPLGRIG--FVNVNHSVLTITLFTMSLATMLPGTVGGYFAVYVWLLGAGAVML 59
Db 8 KGAIRAGSLFGIPFIDVSM--FLILAFVWSYGSGALQAHDEWGTLPFWLTGFSALLL 65
Qy 60 LASLALHAAVAVRRAAGSVESVTLMTFGGVTALGSGAKTPKAAFLIAPATSLAL 119
Db 66 PASVLHLHGSFPAIAGIRVQGISLTFGGVADIERESRSPKQALVVALAGPLVSLAL 125
Qy 120 SATGAL--ATTLGAVTRPA--IVISAVMLATVNLGLFNLLPGAPLDGRLVAYLMR 176
Db 126 FGLFGLGSLALSLAG--PAGAVSL--LAWNLALAI FNNLPGILPDGGVNLKALVMG 179
Qy 177 RHGDSVRAGIGAAARGVVALVLTALGLAEFPVAGLVGGVWLAFLGWIPIPAAREEETRI 236
Db 180 VTGNQYKIRFAGITGGGAGVLMGLG--FVIGNF--NGLMFAIIGWFPFSNARYSEVA 236
Qy 237 STQQLFACVAVDAMTAQPHRTAPGWINVEDFIQRVILGRHAGYPAVDADGSLTGLV-- 293
Db 237 RVQGLSLGTLVQAQVTRTEPAVPAHASLSRFADLYALSSQVSLVTDYDGLVVRIDRG 296
Qy 294 -ALRLRLVAPSRSTTSVGDIALPLHSVTPARQOEPLTALBERMAPLGRSRALVTESG 351
Db 297 ALLRQPTET---WATTPSAVMQGVDASETVAADQPLDEVLTIRLQEKRLPIPIVLPQPSG 352
Qy 352 SAVVGIVTPSDVARLIDYTRLAQPEP 377

Db 353 -LLVGVRIEDIQRF-----FARKEP 372

RESULT 11
O9HNB6_HALSA PRELIMINARY; PRT; 390 AA.
ID O9HNB6_HALSA
AC O9HNB6;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Vng2168C.
GN OrderedLocusNames=VNG2168C;
OS Halobacterium salinarum (Halobacterium halobium).
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
OC NCBI_TaxID=2242;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=NRC-1 / ATCC 700922 / JCM 11081;
RX MEDLINE=20504483; PubMed=11016950; DOI=10.1073/pnas.190337797;
RA Ng W.V., Kennedy S.P., Mahaias G.G., Bergquist B., Pan M.,
RA Shukla H.D., Laezy S.R., Baliga N.S., Thorsson V., Sbrogna J.,
RA Swartzell S., Weir D., Hall J., Dahl T.A., Weiti R., Goo Y.A.,
RA Leithausen B., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA Maddocks D.G., Jablonksi P.E., Krebs M.P., Angevine C.M., Dale H.,
RA Isenbarger T.A., Peck R.F., Pohlschroder M., Spudich J.L., Jung K.-H.,
RA Alam M., Freltas T., Hou S., Daniele C.J., Dennis P.P., Omer A.D.,
RA Ehardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;
RT "Genome sequence of Halobacterium species NRC-1.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181 (2000).
DR EMBL; AE005104; MAG20304.1; -; Genomic DNA.
DR PIR; D84367; D84367.
DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR00644; CBS.
DR InterPro; IPR008915; Peptidase_M50.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR Pfam; PF00571; CBS; 1.
DR Pfam; PF02163; Peptidase_M50; 1.
DR SMART; SM00116; CBS; 2.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
KW Complete proteome.
KW SEQUENCE 390 AA; 41033 MW; BEF392BEB8B8F0CE CRC64;

Query Match 23.1%; Score 459; DB 2; Length 390;
Best Local Similarity 33.9%; Pred. No. 1,1e-23;
Matches 132; Conservative 63; Mismatches 152; Indels 42; Gaps 14;

Qy 7 LGRAGFVNVNHSVLTITLFTMSLATMLP--GTVGGYV-----AVYVWML 51
Db 6 IGRVFGIPKIDVTFLLVLPVFAYLIGTQLEWVSTINGAFPAAGLDATALTAGSVRWL 65
Qy 52 GAGGAVMLASLALHAAVAVRRAAGSVESVTLMTFGGVTALGSGAKTPKAAFLIAPFA 111
Db 66 GAAAVGLFPGVVALHEHGSVVAARKYDFTIDATILMWGGVASITDPBEBNDQSLALAL 125
Qy 112 GPATSLASA--TFGALATITAGVTRPAIVISAVMLATVNLGLFNLLPGAPLDGRLV 170
Db 126 GPASVVALAAVAVGLVALPESLDTVRFVG--YALMLNLALAAFNLLPGFPMDGGRVL 182
Qy 171 RAYIMRRHGSVRAGIGAAARGVVALVLTALGLAEFPVAGLVGGVWLAFLGWIPIPAAR 230
Db 183 RALLARTR--TPARATKIAAEVGLFAFVLGIAGLSF-----NVLIGAFPIYIGAA 234
Qy 231 EETRISTQQLFACVAVDAMT--AOPHTAPGWINVEDFIQRVILGRHAGYPAVDADGSG 288
Db 235 GEAQRTANNAAFEGVTYADVWTPASDVHTVAATASVADLMS--MLEHRHTYPIF--RDAT 292
Qy 289 ITGLVALRLDVAAPSRSTTSVGDIALPLHSVTPARQOEPLTALBERMAPLGRSRA 346
Db 293 AVGWVTLDDASVAVERDARVADVMSDVYTTIPGSADATDALDLOEHSV-----GRL 347
Qy 347 LVTEGSA--VGVITVTPSDVARLIDYTRLAQ 374

```
Db      348 LVVDADGEMVGLTRSD---LMDAFGLIQ 373

RESULT 12
ID      Q64D13_9ARCH PRELIMINARY; PRT; 368 AA.
AC      Q64D13;
DT      25-OCT-2004 (TREMBLrel. 28, Created)
DT      25-OCT-2004 (TREMBLrel. 28, Last sequence update)
DT      25-OCT-2004 (TREMBLrel. 28, Last annotation update)
DE      Zn-dependent proteases.
GN      ORFNames=G218C8_21;
OS      Uncultured archaeon GZfos18C8.
OC      Archaea; environmental samples.
OX      NCBI_TaxID=285411;
RN      [1]
RP      NUCLEOTIDE SEQUENCE.
RX      PubMed=15353801; DOI=10.1126/science.1100025;
RA      Hallam S.J., Putnam N., Preston C.M., Dettler J.C., Rokhsar D.,
RA      Richardson P.M., Delong E.F.;
RT      "Reverse methanogenesis: testing the hypothesis with environmental
RT      genomes."
RL      Science 305:1457-1462 (2004).
RN      [2]
RP      NUCLEOTIDE SEQUENCE.
RA      Putnam N., Dettler J.C., Richardson P.M., Rokhsar D.;
RL      Submitted (AUG-2004) to the EMBL/Genbank/DBJ databases.
DR      EMBL, AY714825; AU082544.1; -; Genomic_DNA.
DR      GO; GO:0004222; F:metalloendopeptidase activity; IEA.
DR      GO; GO:0008233; F:peptidase activity; IEA.
DR      GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR      InterPro; IPR000644; CBS.
DR      InterPro; IPR008915; Peptidase_M50.
DR      InterPro; IPR006025; Pept_M_Zn_BS.
DR      Pfam; PF00571; CBS; 1.
DR      Pfam; PF02163; Peptidase_M50; 1.
DR      SMART; SM00116; CBS; 2.
DR      PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
KW      Protease.
SQ      SEQUENCE 368 AA; 40972 MW; 76ED9972F073F3C CRC64;

Query Match      21.5%; Score 427; DB 2; Length 368;
Best Local Similarity 29.8%; Pred. No. 1.6e-21;
Matches 114; Conservative 76; Mismatches 165; Indels 28; Gaps 10;

Qy      1 MEDAIPGRIAGFVNVNMSVIVITMLFTWSLATMLPGVGYIPAV---VTMLGAGGA 56
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1 MKTSIQIKVWGIPKIHISFLVLVLPVVFAN--DAVFQADVAVSLTFLVTLTLT 57
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy      57 VMLASLAAHELAAVAVRAGVSVESTLMFGVTLALGGEAKTPKAFRIAPAPATS 116
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      58 ILLPSCVLLHGHGSHWAVRIGIRSTILLIGIAAEEVPRDPRAEKRISIGPLVS 117
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy      117 LALS-----ATFGALATITLAVGRTPAIVISVAMWMLATVNLGLFNLLFGAPLDGRLVR 171
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      118 LTIGVLCVLAIVGLGIYDMVQVSTISHFLMSIAYINIVAPFNLLIPAFMDGRLVR 177
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy      172 AYLRRHGDSTVAGIGAAPAGRVALLVIALGLAEFVAGLVGWLAIFGFIIPAARE 231
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      178 AW-YAGHPHYLAKTKAVHIGMFAIVMGVL-----GLPASLWLLIAFIYIGASE 228
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy      222 EETRISTQGLFAGVAVADMTAOPHTAPGMINVEDFIORVYVGERSAVPVADR-DGSIT 290
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      229 EEKTEVSTVTEGIRVRLMTREIRIAYVQDNLTISEL-RLMEKKGHLGIPVVDQGTGKI 287
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy      291 GLVALROLADVAPSRSTTSVGDIALPLHSVPTAPQPEBLTALLERMAPLGRSPALYTE 350
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      288 GIVTFDIRSVPMSEHGVLVREV-MAKVVVFIPEDADAMAL--KMSTENWVGLVQD 344
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy      351 GSAVVGIYTPSPVAKLIDV--YR 371
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      345 RGAITGIIVSRITLTRSIETLGYR 367
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

```
RESULT 13
ID      Y528 SYN3 STANDARD; PRT; 379 AA.
AC      O5518;
DT      01-NOV-1997 (Rel. 35, Created)
DT      01-NOV-1997 (Rel. 35, Last sequence update)
DT      13-SEP-2005 (Rel. 48, Last annotation update)
DE      Hypothetical protein gl10528 (EC 3.4.24.-).
GN      OrderedLocustNames=gl10528;
OS      Synchocystis sp. (strain PCC 6803).
OC      Bacteria; Cyanobacteria; Chroococcales; Synchocystis.
OX      NCBI_TaxID=1148;
RN      [1]
RP      NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX      MEDLINE=96127529; PubMed=8590279;
RA      Kaneko T., Tanaka A., Sato S., Kotani H., Sazuka T., Miyajima N.,
RA      Sugiyura M., Tabata S.;
RT      "Sequence analysis of the genome of the unicellular cyanobacterium
RT      Synchocystis sp. strain PCC6803. I. Sequence features in the 1 Mb
RT      region from map positions 64% to 92% of the genome."
RL      DNA Res. 2:153-166 (1995).
RN      [2]
RP      NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX      MEDLINE=97061201; PubMed=8905231;
RA      Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA      Miyajima N., Hirosewa M., Sugiyura M., Sasamoto S., Kimura T.,
RA      Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
RA      Shimpou S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yaouda M.,
RA      Tabata S.;
RT      "Sequence analysis of the genome of the unicellular cyanobacterium
RT      Synchocystis sp. strain PCC6803. II. Sequence determination of the
RT      entire genome and assignment of potential protein-coding regions."
RL      DNA Res. 3:109-136 (1996).
CC      -1- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC      -1- SUBCELLULAR LOCATION: Integral membrane protein (potential).
CC      -1- SIMILARITY: Belongs to the peptidase M50B family.
CC      -1- SIMILARITY: Contains 2 CBS domains.
CC      -----
CC      This Swiss-Prot entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use as long as its content is in no way modified and this statement is not
CC      removed.
CC      -----
DR      EMBL; BA000022; BAA10876.1; -; Genomic_DNA.
DR      PIR; S76029; S76029.
DR      InterPro; IPR000644; CBS.
DR      InterPro; IPR006025; Pept_M_Zn_BS.
DR      InterPro; IPR008915; Peptidase_M50.
DR      Pfam; PF00571; CBS; 1.
DR      Pfam; PF02163; Peptidase_M50; 1.
DR      SMART; SM00116; CBS; 2.
DR      PROSITE; PS00142; ZINC_PROTEASE; 1.
KW      CBS domain; Complete proteome; Hydrolase; Hypothetical protein;
KW      Metal-binding; Metalloprotease; Protease; Repeat; Transmembrane; zinc.
FT      TRANSMEM 20
FT      TRANSMEM 40
FT      TRANSMEM 54
FT      TRANSMEM 115
FT      TRANSMEM 148
FT      TRANSMEM 212
FT      TRANSMEM 260
FT      TRANSMEM 310
FT      TRANSMEM 327
FT      TRANSMEM 375
FT      ACT_SITE 76
FT      ACT_SITE 77
FT      METAL 75
FT      METAL 79
FT      METAL 79
SQ      SEQUENCE 379 AA; 40465 MW; 1CC3251650078ECB CRC64;

Query Match      21.4%; Score 425.5; DB 1; Length 379;
Best Local Similarity 30.0%; Pred. No. 2.2e-21;
Matches 115; Conservative 70; Mismatches 147; Indels 51; Gaps 11;
```

QY 1 MRDAIPGIRING--FVNVNHSVLTMLFTMSLATML---PGTVGYPAVVYMLLGAG 55
DB 11 MNNNIRVSGSLFGIPVNPSPW--FLLGLVLTLSYGQDARFPQSGGTP---WLLGLIT 64
QY 56 AVMLLASLAEHLAAVAVARRAGVSVESVTLMLFGCVATLAGEAATPKAAFRIPAAGPAT 115
DB 65 ALLLPASVVAHELGSLVALAGIVKSTLFLFGGLASLEKESNTPMQAFVAVLAGNAV 124
QY 116 SLASATFGALATITAGVTPPAIV--ISVAMVLTAVNLLGLFNLLPGAPLDGGRLVAV 173
DB 125 SLVL-----FLGLITVGTQIPLPVPQQAIIIGLGMINLALFNLI PGLPLDGGVNLKSI 179
QY 174 LMRHSDSVRAGIGAAAGRVAVLIALGLAEFVAGLVGVGWLAFICGMPFAAREE 233
DB 180 VMQITGNQNKGLIASRVGGQGMALAIIGISIGLITLIPGSEFTLLIGWFLQNASSA 239
QY 234 TRISTQCFAGRVADMTAQTAPGMINVEDFIQRYVLGER--SAVPAVDRDSITGL 292
DB 240 RNAQVKEQMEAFVADAVIPIPSPIIPAGLNIREFANDVYIGKTPRRRLVIGADNQLGV 299
QY 293 VALRQLRDVAPSRSTSVGDIALPLHSVPTAR--POEPLTALLERMAPLGRSRALVTEG 351
DB 300 LATEDIK-----HVPTSDMPQVTVDSIMQY-----PQWVTVAN 334
QY 352 SAVVGIVTPSDVARIIDVYRLAQ 374
DB 335 QSLF-----EVAQLIDQOKLSE 351

RESULT 14
Q7VE60_PROMM PRELIMINARY; PRT; 422 AA.

AC Q7VE60; 01-OCT-2003 (TREMBLrel. 25, Created)
DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Hypothetical protein.
GN OrderedLocustNames=PW1225;
OS Prochlorococcus marinus (strain MIT 9313).
OC Bacteria; Cyanobacteria; Prochlorales; Prochlorococaceae;
OC Prochlorococcus.
OX NCBI_TaxID=74547;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22825698; PubMed=12917642; DOI=10.1038/nature01947;
RA Rocap G., Larimer F.W., Lamerdin J.E., Malfatti S., Chain P.,
RA Ahlgren N.A., Arellano A., Coleman M., Hauser R., Hess W.R.,
RA Johnson Z.I., Land M.L., Lindell D., Post A.F., Regala W., Shah M.,
RA Shaw S.L., Steglich C., Sullivan M.B., Ting C.S., Tolonen A.,
RA Webb E.A., Zinser E.R., Chisholm S.W.;
RA "Genome divergence in two Prochlorococcus ecotypes reflects oceanic
RT niche differentiation."
RL Nature 424:1042-1047(2003).
EMBL: BX572098; CAZ1400.1; -: Genomic_DNA.
DR GO: GO:0004222; F:metalloendopeptidase activity; IEA.
DR GO: GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro: IPR006915; Peptidase_M50.
DR InterPro: IPR006025; Pept_M_Zn_BS.
DR Pfam: PF02163; Peptidase_M50; 1.
DR PROSITE: PS00142; ZINC_PROTEASE; UNKNOWN_1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 422 AA; 45664 MW; 7C0E619FDD1F948 CRC64;

Query Match 21.3%; Score 424; DB 2; Length 422;
Best Local Similarity 29.5%; Pred. No. 3, 1e-21;
Matches 127; Conservative 69; Mismatches 177; Indels 58; Gaps 10;

QY 1 MRDAIPGIRINGFVNVNHSVLTMLFTMSLATMLPGTVGYPAVVYMLLGAGAVML 59
DB 1 MGEQWELMKIRIGIPLRVHPSWVFIILLFTWISQNVSAASSLPAMWSGIGLITALL 60
QY 60 LASLAEHLAAVAVARRAGVSVESVTLMLFGCVATLAGEAATPKAAFRIPAAGPATSLAL 119

DB 61 FLVLAEHLGSLVALAEQVKVRSITLFLGCVASVERECSTPMASLRVAAGPLVSLVL 120
QY 120 SATFGALATITAGY---RTPAIVTSVAMVLTAVNLLGLFNLLPGAPLDGGRLVAVYLW 175
DB 121 -----AAVALLTGVAADHVPPLANTVGOIGLGNLILLALFNLLPGAPLDGGRLTKALW 175
QY 176 RRHDSVRAGIGAAAR--GRVAVLIAL--GLAEFVAGLVGVGWLAFICGMPFAAREE 233
DB 176 QWTS--SQRKGVQVATVAGRALSTLSAMVGLGLFVKGGIGGLMLLMGFGAGASHSQ 234
QY 234 TRISTQCFAGRVADMTAQTAPGMINVEDFIQRYVLGERHSA---YPAVDRDSIT 290
DB 235 QLLALQVLAELNANGQAAGNFRLVEDDQSLRSLQRLSGSERQSGEWLVGRSRKW 294
QY 291 GLVALRQLRDVAPSRSTSVGDIALPLHSVPTARPOEPL----- 330
DB 295 GYMTDQLEKELPVQWQROCLADHMKRISELPALGERKAPMQVMALEQAEGRLLVFNV 354
QY 331 -----TALLERMAPLGR-----SRALVTEGSNAVVGIVTPSDVARIIDVYR 371
DB 335 AGLPQGLDRIDSEAVLKR---LGVRLPVQPLEAARQNTYPLGMALPVVESVSGGL 411
QY 372 LAQPEPTTS 382
DB 412 VEQPEASSSTS 422

RESULT 15
O29915_ARCFU PRELIMINARY; PRT; 362 AA.

AC O29915; 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Hypothetical protein AF0332.
GN OrderedLocustNames=AF0332;
OS Archaeoglobus fulgidus.
OC Archaea; Buryarchaeota; Archaeoglobi; Archaeoglobales;
OC Archaeoglobaceae; Archaeoglobus.
OX NCBI_TaxID=2234;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=96049343; PubMed=9389475; DOI=10.1038/37052;
RA Kleink H.-P., Clayton R.A., Tomb J.-F., Hickey O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Gwin M.L., White E.K., Peterson J.D.,
RA Richardson D.L., Kervazenga A.R., Graham D.E., Kypides N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.R.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.J.,
RA Peterson S.N., Reich C.I., McNeil L.K., Badger J.H., Glodek A.,
RA Zhou L., Overbeek R., Gocayne J.D., Weidman J.F., McDonald L.A.,
RA Usterback T.R., Cotton M.D., Spriggs T., Artlich P., Kaine B.P.,
RA Sykes S.M., Sadow P.W., D'Andrea K.P., Bowman C., Fujii C.,
RA Garland S.A., Mason T.M., Olsen G.J., Fraser C.M., Smith H.O.,
RA Moese C.R., Venter J.C.;
RA "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus."
RL Nature 390:364-370(1997).
EMBL: AE001082; AB90905.1; -: Genomic_DNA.
DR PIR: D69291; D69291.
DR TIGR: AF0332; -;
DR GO: GO:0004222; F:metalloendopeptidase activity; IEA.
DR GO: GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro: IPR000644; CBS.
DR InterPro: IPR008915; Peptidase_M50.
DR InterPro: IPR006025; Pept_M_Zn_BS.
DR Pfam: PF00571; CBS; 1.
DR Pfam: PF02163; Peptidase_M50; 1.
DR SMART: SM00116; CBS; 2.
DR PROSITE: PS00142; ZINC_PROTEASE; UNKNOWN_1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 362 AA; 39865 MW; A09CF565765D7662 CRC64;

Query Match	21.0%;	Score 418;	DB 2;	Length 362;
Best Local Similarity	30.5%;	Pred. No. 6.8e-21;		
Matches 113;	Conservative 70;	Mismatches 162;	Indels 26;	Gaps 10;

Qy	MPRAIPGRAGVAVNWHMSVLVILMLFTWLSLA-TMLPGTGGVPAVYVYMLLAGGAVML	59
Ds	1 MRASFKIRFVGGIDVEVHISLILVLVLIYASVSPPFYGRANFPLSRRTILSSMAAVGL	60
Qy	60 LASLLAHLHAHAVVARAGVSVEVYTLMLFGVTALGSEAKTPKAFAPIAAGPATSLAL	119
Ds	61 FASILAHLEGHSLVARVGRIRGMILFIFGVVANMDELPPKPRELVALISGPAIS---	117
Qy	120 SATFGALATTLAGVRRPAIVSYAMWL-ATVWLLIGLFNLLPGAPLDGGRLVRYLIMRRH	178
Ds	118 ---FGIAVVSALLSSIPVAELSAFMLFEGYLMFLIATFNLLPAPFMODGRILRSFLAEKR	174
Qy	179 GDSVPAIGIGARAGRVRVALVLTALGLAEPAVGWGLVFIQGFIPAAAREETRIST	238
Ds	175 SYVEAKTK-AAEIGRLALIFMAIF-----GIFTNMILLIALFYITGANEERVLVL	225
Qy	239 QQLFAGVRVADAMTAOPHTAPGMINVEDEFIQRYVVGEBHSAPVADRDGSITGLVALROL	298
Ds	226 ENVLGRVRADVNMTEVVTPEMTESEVID-LIKTTHLGEPPVVEGS-RLVGIITLHDI	283
Qy	299 RDVAPSRRTSTSGDIALPLHSYPTPAKPOEPTALLERMAPLGGPSRALVTESGAVYGIY	358
Ds	284 IGVPEPER-----VGNIM--SREVVAVSPQSAFEAFKIMSEWG-IGRLPVEHGRVVGIV	336
Qy	359 TPDSVAVRLIDV	369
Ds	337 SRSDLMRAIKETI	347

Search completed: March 23, 2006, 05:24:28
Job time : 102.515 secs

This Page Blank (uspto)

Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 13409
; LENGTH: 346
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-13409

Query Match 8.7%; Score 174; DB 2; Length 346;
Best Local Similarity 27.4%; Pred. No. 4,3e-09;
Matches 92; Conservative 37; Mismatches 137; Indels 70; Gaps 17;

QY 58 MLASLALHAEIAVAVARRAGVSVTLMFGVTALGGEAK-TPKAAFRIPAPGPTS 116
DB 19 IVPASVILHEIGHALAFRRYGGPAR-IELHGMGGTQTHDAHLTHQSAFVSFAPGIG 77
QY 117 LALSFGALA-----ITLAGVTRPAIVISVAMWLATVNLGLFNLLPGAPDGG 166
DB 78 FLAGGLIGLSQVLPGSPGLADGVR-----QFLM-----VNIQKMLFNLLPQPLDG 127
QY 167 GR---LVRAYLWRRHGSVRAGIGARVVALVLLALGAEFVAGLVGVWLAFIG 222
DB 128 GHLLADLYRARSQYHERGV-LGVG-----IATAVVLGLAIW-SKQLMWGMAMVIG 178
QY 223 WFTPAABEERTISTQQLFAGVRAVADMTAOPHTAP-GWIVNEFIRYVLGERSAYP 281
DB 179 VWNLEQLRTPRRAPERRFALPR---LVKPEAPAGALISEQLMDEL---RGTRP 230
QY 282 --VADRDSITG---LVALRQLRDVAPSRSTSVGDIALPLHSVPTARQEPPL--- 330
DB 231 PGADADDDLEGPDPALVGEMLDNGIPE---LAVSLSQAFQAPLARTGHALVIAL 286
QY 331 --TALLERMAPLGPSSR-----ALVTEGSAV 355
DB 287 LHTGRLGELASLDBSSSHARQISEPTLALISEHAGTV 322

RESULT 3
US-09-252-991A-29289
; Sequence 29289, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 29289
; LENGTH: 367
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-29289

Query Match 7.0%; Score 139.5; DB 2; Length 367;
Best Local Similarity 24.9%; Pred. No. 1.5e-05;
Matches 98; Conservative 48; Mismatches 143; Indels 105; Gaps 18;

QY 24 ILMLFTSLATMLPGTVGIPAVVY-----WLIGAGAVMLASLAEIAHAAVARRAG 78
DB 27 VTMRLYW--ATPLVAGLGGWLASLVGWPPLPMITGSLAVMLVRLADLPLAEVPGARKCG 84
QY 79 VSVSIVTLMLFG-CVTLALGGEAKTPKAAFRIPAPGPTSIALSA--TFGALATLAVRT 135
DB 85 -----QWIVIGDI-----GLHFPVAVIEQV-----LAHSVIVFGAVATLSSVLA 125
QY 136 PAIVISVAMWLATVNLGLFNLLPGAPLDGRLVRAVYWRHDSVRAGIGARRVV 195
DB 126 IAFMRSGEDBATAF-----ASMFGASSEMNLGGRHG-AVLSVAAAQSLRL 174
QY 196 ALVIALGLAEFVAGLVGV-----WLA-----IGFTPAAREEFT-- 235
DB 175 LVVLVLPAAEQYLLGGQPGPHQAAPVDMWMLALLFPAGALVALGW-----QKLRP 226
QY 236 ----ISTQQLPAGRVA-DAMTAQPHAPG---WINVEDFIQRVYVGBERRSAYPVARDG 287
DB 227 NPMLLGPULLAAGVSLGPDHLHIGLPASSGVGWL-----IG 263
QY 288 SITGLVALRQLRDVAPSRSTSVGDIALPLHSVPTARQEPPLT-----ALLERMAPLGP 342
DB 264 SALGCHNRSPFRAPAPVERTIVCTLMMFPAALAAELGMLTTLDHOSIMLGMRGCI 323
QY 343 RSRLVTEGSAV-VGIPTSDVARLIDVYRLAQP 375
DB 324 AELSLTAEALQLSVPLVVALQVLRLLLVFLAEP 357

RESULT 4
US-09-712-363-257
; Sequence 257, Application US/09712363
; Patent No. 6892139
; GENERAL INFORMATION:
; APPLICANT: Eisenberg, David
; APPLICANT: Rochelein, Sergio H.
; APPLICANT: Marcotte, Edward M.
; TITLE OF INVENTION: DETERMINING THE FUNCTIONS AND
; FILE REFERENCE: 07419-032001
; CURRENT APPLICATION NUMBER: US/09/712,363
; PRIOR FILING DATE: 2000-11-13
; PRIOR APPLICATION NUMBER: PCT/US00/02246
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/179,531
; PRIOR FILING DATE: 2000-02-01
; PRIOR APPLICATION NUMBER: 60/117,844
; PRIOR FILING DATE: 1999-01-29
; PRIOR APPLICATION NUMBER: 60/118,206
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: 60/126,593
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 60/134,093
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 60/134,092
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 60/165,124
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/165,086
; PRIOR FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 292
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 257
; LENGTH: 404
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-09-712-363-257

Query Match 6.3%; Score 125; DB 2; Length 404;
Best Local Similarity 28.9%; Pred. No. 0.00051;
Matches 57; Conservative 20; Mismatches 66; Indels 54; Gaps 9;

NUMBER OF SEQ ID NOS: 16625
; SEQ ID NO 15926
; LENGTH: 472
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-15926

Query Match 6.1%; Score 120.5; DB 2; Length 472;
Best Local Similarity 22.9%; Pred. No. 0.0019;
Matches 92; Conservative 53; Mismatches 133; Indels 123; Gaps 21;

QY 38 GTVGVPVAVVWMLGAGGAVMLASL--AHELA-----HAAVARRAGVSVESTL 86
DB 107 GSMIGHPAV--TGLAVFLLLATSMFGAFELAPSSFQTKLVAGG--AGVA---GA 157
QY 87 WLFPGVTA-LGGEAKTPRAAFRIAPAGPATSLATSGALATITLAVRTPAIVISV--- 142
DB 158 FLMSVSGFLAPCTGCVLTGTLAVAKSANTTLGASL--LFVYALGIVFPFLIGVTV 215
QY 143 ----AMWLATVNLILGLFNLLPGAPLDGRLVRAVYLRHRGDSVRA-GIGAAAR-GRV 195
DB 216 RLPRGVMMWVKSVTGIIMLV---ALAFSYLKDAFPWAR--DVYKGLGLHVGRVPGAVI 269
QY 196 ALVLIAGLA-----EF-----VAGGLVGVWLAFIG 222
DB 270 AALLVAVGVALGAVHRSEKGAARESPALGVALVVALVIRGALDGGPVTLMVS-LG 328
QY 223 WFIPIAAAREESTRISTOOLFAGVAVADAMTAQPTAPGM-----INVEDFIOR 270
DB 329 -----LAEPAAPAMQHHWPAKQATPSPEAPDQ- 358
QY 271 VYIGERSAVVADRDSITGLVALRQI-RDVAERSSTSVGDIALPLHSVPTARPOEP 329
DB 359 -VLAQAKREGRPVLIDFPADWCAKELDRDTPYPAQVVISSEDEGRFINIKIDATNSDS 417
QY 330 LTALLERMAPIGPRSRALVT-EGSA-----VVGITPBSDVA 364
DB 418 LDALLERGVEGLPVAFISPEGKVLTPQPRVTFGLAPSPFA 458

RESULT 8

US-09-252-991A-19506
; Sequence 19506, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19506
; LENGTH: 691
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19506

Query Match 6.1%; Score 120.5; DB 2; Length 691;
Best Local Similarity 25.7%; Pred. No. 0.0033;
Matches 106; Conservative 40; Mismatches 166; Indels 101; Gaps 20;

QY 50 LLAGGAVMLASLAEHLAAVAVARRA--GVSVESVTLMFGVTTALGSEAKTP----- 102
DB 78 VLAHDHALGVVALGTGELGALRLVPGIGQLGAVAGLHGAN--VGGTQPAFLRVLV 135
QY 103 -KAAPRIAPAGPATSLATSGALATITLAVR-----TPAIVISVA---MWLAVTNLL 153
DB 136 GRVEDRVGIALPVVA-AVPARGGOATTFLLAVRIASGDLPADVLFVGHSDVQRHANLAG 194

QY 154 GLFNLLPGAPLDGRLVRAVYLRHRGDSVRAGIG---AARAGVVALVLIAGLAEFVA 209
DB 195 GVEGLIDGL---GPROVEAVNOGLVGLPDAPVQSVALLQRORLAEVLVLAQGAVFQRA 251
QY 210 -GGLVGV-----WLAFIGFIFPAAREESTRISTOOLFAGVAVADAMTAQPH- 256
DB 252 MAGFVDLAPRLALDEGFLAFAGMLPLAVGVITEVALASQLF---QVEILIVAQDHG 308
QY 257 TAPMINVEDFIORYVYIGERSAVPVADRDSITGLVALRQLRDVARSSTTS---VG 312
DB 309 HAPQLAVE-----ABHHARQTS--DGNARGVLGRADLHVPHRRHAQRQGVVYG 357
QY 313 DIA-----PLHSVPTARPOEPTLRLERMAPIGPRSRAL--VTEGSA 353
DB 358 HQALAAALRRDSPVVRGDDPEHAORCDLPRRLDREAGNLPVQAQALEPRLVERQV 417
QY 354 VVGI-----VTESDVAR-----LIDVYRLAQPE-----PFTT 381
DB 418 FVGIQRQPGQLVAADLRLERQRGDPFLQVHRQAEVQEGEDQGRVLDLPVFR 470

RESULT 9

US-09-489-039A-7541
; Sequence 7541, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 7541
; LENGTH: 508
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-7541

Query Match 6.0%; Score 120; DB 2; Length 508;
Best Local Similarity 23.6%; Pred. No. 0.0023;
Matches 86; Conservative 47; Mismatches 140; Indels 92; Gaps 18;

QY 10 IAGFVAVWVH--SVYVILMLFTWSLATMLP-GTVGGVP----- 44
DB 183 VGGFLA-VSMTDVTYQASLMIFALITLTPVIYIISVGGGDSLEVIKOKSIENIDMLKGLNF 241
QY 45 ----AVVYMLLAGGAVMLASLAEHLAAVAV-ARRAGVSVESTLML--FGVTALG 96
DB 242 VAISLWGMGLGYFGQCHIIILAFMAADSHSHSVHARRISMT-----WMILCIGGAVAVG 295
QY 97 GEATTPRAAFRIAPAGPATSLATSGALAI-----TLAGVTPAIVISVAVW 145
DB 296 -----FFGIAFPNNPISLAGVNONARVFTEIAQLIFNPWIAIGLISALIAV--- 344
QY 146 LATVNLILGLFNLLPGAPLDGRLVRAVYLRHRGDSVRAGIGAAARAGVVALVLIAGL- 204
DB 345 MSTISC-----QLIVCSATITBDLYKFLKKNQOKELVWVGMMV-LVALVALAALAN 398
QY 205 AEFVAGLVGVWLAFIGWF---IFAAREESTRISTOOLFAGVAVADAMTAQPTAPG 260
DB 399 PENRVLGLVAVAMGFAAGFPVVLFSVMSRMTR---NGALAGM-VIGALTIVVMKQFG 454
QY 261 WINVEDFIORYVYIGERSAVPVADRDSITGLVALRQLRDVARSSTSVGDIALPLHS 320
DB 455 WLGXEIIPGVFVS-----IGIVVF-SLDKAPASMQORFAEDAHYHT 499
QY 321 VPTAR 325
DB 500 PEPVR 504

RESULT 10
US-09-902-540-10370
; Sequence 10370, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 10370
; LENGTH: 574
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-10370

Query Match 6.0%; Score 120; DB 2; Length 574;
Best Local Similarity 22.2%; Pred. No. 0.0028;
Matches 95; Conservative 53; Mismatches 126; Indels 154; Gaps 23;

QY 11 AGFVVNVH--WSVLV--ILMLFTSLATMLPCTVG--GYPAVVVYLLGAGCAVMLLAALL 64
DB 204 AGALVSPTGWSMLGAVLTGFLAPAVWSQALCEVTYRALNSWMTGSAVLVSSGLL 263
QY 65 AHELAHAVARAGVSVESVTLMFGVTALGGEA-----KTPKAFRIAPAGPATSLA 118
DB 264 SPAPQMSVAR----SFKSLG--GLFKGKAKEDDPLAGIECPFAWFPLEG----- 310
QY 119 LSATFGALATITAGVTPAIVISVAMWLATVNLGLFNLPGADLDGRLVRAVLMRRH 178
DB 311 --AIGPVAVFELM-----AYLFQIPMWAGVLM-----PL----- 338
QY 179 GDSVRAGIGAR--AGRVVALVILAG--LAEPVAGCL-----VGGVWLAFI- 221
DB 339 --AVMGVIARVGTETDTPTKALGPVTQLFGSLAPGNI PANVMSANATGGVLSHAD 396
QY 222 -----GMFLFAAREESTRISTQOLF--AGVAV-----ADAMTAQPTAPG 260
DB 397 LITDLKSGMLGANRQ-----FVAQLFGVAGAAVVPVKIIVPDASMLGTEEPAPA 452
QY 261 --WINVEDTQRYVYLSRHSAYPADRGSTGLVALRLQLDVAPSRRTSTVGDIAP 317
DB 453 SMVAVGVSKLATGV-----ALPESARWGLCG-----AALG 485
QY 318 LHSVPTARPOEPLTLLERMAPLGRS-----RALVTEGSAVVGIVTPSDVARIID 368
DB 486 IFLV-----LLERMAPAKAKAVVPBPAGFGIAIVIPGSSSIAFPGSSAIAELL- 533
QY 369 VYRLAQPE 376
DB 534 --RRTKPK 539

RESULT 11
US-09-902-540-11981
; Sequence 11981, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540

; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 11981
; LENGTH: 886
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-11981

Query Match 6.0%; Score 120; DB 2; Length 886;
Best Local Similarity 23.0%; Pred. No. 0.0054;
Matches 101; Conservative 44; Mismatches 149; Indels 145; Gaps 24;

QY 46 VVYMLGAGAVMLASL-----LAHELAHAVA-----RRAGVSVESVT 85
DB 235 VAVYVLSAGAVVAGLAFSSRPGDFQVDLASALAPLIAAGVFLLEBRRLVHLMMT 294
QY 86 L-----WLFPGVTALG-----EAKTPKAFRI--AFAPATSLA 118
DB 295 TILAGALFARTQAPPTDPGWWVFGGTVAAGIMPLARWSAFPGLRIRLAWTLTGLSVAMP 354
QY 119 LSATFGALATITAGVTPAIVISV-----W-----LATVNLGLFNLPGAP-LDG 166
DB 355 LVSVQGSASBPQVLTGLVITAGAAHAGKMRWGLHGLFGLVAFVEGTPALEG 414
QY 167 GRLVRAVLMRRHGDVSRAGIGARGRVVALVILAGLAEPVAGLVGVWLAFIWFIF 226
DB 415 S-----W-----GRLAVPALVS--GL--YGVAGLVGAMA----- 440
QY 227 AAAREESTRISTQOLF--FAGRVADAMTAQPTAPGMINVEDTQRYVLSRHSANPVA 283
DB 441 OPASRRDGLPLDLDLALCLAAAGVILAVDSSP--SAPDLVSVTGLASVLT-----SALPFA 495
QY 284 -----DRDS--ITGLVA-----LRQLRDVAPSRRTSTVGDIAPLSHV 321
DB 496 VVTAMLLRVRDSSRLVGFLLASGLAAVSVQVGTSDFA--SPRAALVAASLALGFAPF 554
QY 322 PTAPPOEPLTAL-----LERM-----APLGRSRALVTEGSAVVGIVTPSDVARIIDV 369
DB 555 SALAGRTPLDAADPAAGRRRMFDRLPLGARGPLFTDGFATALV--QTVIALTLA 613
QY 370 YRLAQPEPTFTSPQADR 388
DB 614 GMLAR-----PTDAER 624

RESULT 12
US-09-902-540-15291
; Sequence 15291, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 15291
; LENGTH: 332
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-15291

Query Match 6.0%; Score 119; DB 2; Length 332;
Best Local Similarity 26.4%; Pred. No. 0.0016;
Matches 64; Conservative 19; Mismatches 71; Indels 88; Gaps 13;


```
QY 281 PVADRDSITGLVALRLRD-----VAPSRSTTSVGDIALPL-----HSVPT----- 323
DB 1030 -VLGRDGA-AGLDAGQAFADSGFNSLSAVELNRRLTAATVATLPAFAIPDHPTELAQY 1087
QY 324 -----ARPEPLTALLE-----RMAPLGRSR----- 345
DB 1088 LITQIDHGSSAAAAANPARIIDALTDPLQACDAGRADGWMKVALASNTREMSPPVR 1147
QY 346 -----ALVTEG-SAVGVITPS-----DVARLI-----DVYRLAQBEPTFTTS 382
DB 1148 NNVSXKVALADGISDVVICIPTLTVLSDQREYRDIANAMTGRHSVSLTL--REFDSS 1205
QY 383 ---PODAD 387
DB 1206 DALPQNAD 1213
```

RESULT 15

```
US-09-198-452A-653
; Sequence 653, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Gliffais, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
; TITLE OF INVENTION: and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 653
; LENGTH: 551
; TYPE: PRF
; ORGANISM: Chlamydia pneumoniae
US-09-198-452A-653
```

```
Query Match 5.8%; Score 115.5; DB 2; Length 551;
Best Local Similarity 21.4%; Pred. No. 0.0075;
Matches 65; Conservative 40; Mismatches 94; Indels 105; Gaps 14;
```

```
QY 1 MRDAIPLGRIGAFVNVNH-----MSVLVILMLFTWSLATMLP--GTVGGYPA 45
DB 140 LQELPQG-LRGFIWVRYSYSIYYVMSLSVTLMLF-WGLANQITITTEAGRFPYA 197
QY 46 VV-----VYLLGAGAVMLASLAEHLAAVAVARAGVSVE--VTLM 88
DB 198 LINTGLNLSSTICAGEISTWM---GKQTFVAYSFACDSWHSVMLNLTMLTSCGLIMIML 253
QY 89 F-----GGVTALGCEAKTPKAAFRIFAGPATSLASATPGAL 126
DB 254 YRRIHHLITDISIPPSRRLTAEGEAATNLKKEKKPKAKARLNF---LHLQSRVILGL 309
QY 127 AITLAGVTRPAIVISVAV-----WLATVNLILGLFNLPGAPLDGRLV 170
DB 310 AIIIVSYNLVILHLEFVVMQDVQSIIYSHVEFNWMSRITTLIGVVSVL-AAVLLTGQCI 368
QY 171 RAYLWRRHGDSVRAGIGABARVVALVLIAGLAEF-----VAGSLVGGWMLA 219
DB 369 RKMGMF-----VGAL---VTPVLVVSGLLFFGTIFAKRDISIFGGVIGMTPLA 415
QY 220 FIGW 223
DB 416 LAAW 419
```

Search completed: March 23, 2006, 05:30:18
Job time : 28.9039 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using SW model

Run on: March 23, 2006, 06:14:33 ; Search time 82.6454 Seconds
(without alignments)
1986.885 Million cell updates/sec

Title: US-10-617-038-28

Perfect score: 1990
Sequence: 1 MRDA1PLGR1AGFVNVNWS.....QPEPTTTSPODADRFS DAG 393

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA Main:
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1990	100.0	393	4	US-10-617-038-28 Sequence 28, Appl
2	425.5	21.4	379	4	US-10-047-260-12 Sequence 12, Appl
3	311	15.6	376	4	US-10-156-761-14201 Sequence 14201, A
4	143	7.2	473	4	US-10-156-761-14437 Sequence 14437, A
5	140	7.0	591	5	US-10-482-706-183 Sequence 183, App
6	140	7.0	649	4	US-10-080-170-24 Sequence 24, Appl
7	140	7.0	649	4	US-10-080-170-24 Sequence 24, Appl
8	140	7.0	649	4	US-10-468-356-24 Sequence 24, Appl
9	138.5	7.0	706	4	US-10-282-122A-50265 Sequence 50265, A
10	133.5	6.7	570	4	US-10-156-761-9868 Sequence 9868, Ap
11	133	6.7	944	5	US-10-493-462-25 Sequence 25, Appl
12	132	6.6	3745	4	US-10-205-032-14 Sequence 14, Appl
13	130.5	6.6	3362	4	US-10-378-083-6 Sequence 6, Appl
14	130.5	6.6	4685	4	US-10-156-761-10433 Sequence 10433, A
15	128	6.4	487	4	US-10-437-963-17542 Sequence 17542, A
16	127.5	6.4	400	4	US-10-282-122A-48123 Sequence 48123, A
17	126.5	6.4	851	3	US-09-991-936-1915 Sequence 1915, Ap
18	126.5	6.4	851	5	US-10-978-245-1915 Sequence 1915, Ap
19	126.5	6.4	3808	4	US-10-378-083-7 Sequence 7, Appl
20	125.5	6.3	466	4	US-10-156-761-8852 Sequence 8852, Ap
21	125	6.3	404	3	US-09-712-363-257 Sequence 257, App
22	125	6.3	533	4	US-10-424-599-215899 Sequence 215899, A
23	122	6.1	1029	4	US-10-389-647-613 Sequence 613, App
24	121.5	6.1	477	4	US-10-425-115-310459 Sequence 310459, A
25	121	6.1	833	4	US-10-156-761-12366 Sequence 12366, A
26	120.5	6.1	450	4	US-10-156-761-11177 Sequence 11177, A
27	120.5	6.1	457	3	US-09-738-626-4139 Sequence 4139, Ap

28	119	6.0	413	4	US-10-282-122A-62054 Sequence 62054, A
29	119	6.0	811	5	US-10-450-763-59793 Sequence 59793, A
30	118	5.9	791	4	US-10-156-761-12134 Sequence 12134, A
31	117.5	5.9	323	4	US-10-156-761-8143 Sequence 8143, Ap
32	116.5	5.9	260	4	US-10-282-122A-61660 Sequence 61660, A
33	116.5	5.9	531	4	US-10-437-963-158999 Sequence 158999, A
34	116	5.8	1402	3	US-09-712-363-166 Sequence 166, App
35	116	5.8	7257	5	US-10-732-923-20621 Sequence 20621, A
36	115.5	5.8	384	4	US-10-156-761-9976 Sequence 9976, Ap
37	115.5	5.8	385	4	US-10-437-963-167452 Sequence 167452, A
38	115.5	5.8	468	4	US-10-425-114-57349 Sequence 57349, A
39	115.5	5.8	540	4	US-10-282-122A-54741 Sequence 54741, A
40	115.5	5.8	551	4	US-10-289-762-653 Sequence 653, App
41	115.5	5.8	561	4	US-10-425-115-516538 Sequence 516538, A
42	115.5	5.8	1061	4	US-10-282-122A-49840 Sequence 49840, A
43	115	5.8	403	3	US-09-976-059-23 Sequence 23, Appl
44	115	5.8	548	4	US-10-282-122A-62037 Sequence 62037, A
45	115	5.8	596	5	US-10-762-107-82 Sequence 82, Appl

ALIGNMENTS

```
RESULT 1
US-10-617-038-28
; Sequence 28, Application US/10617038
; Publication No. US20040057963A1
; GENERAL INFORMATION:
; APPLICANT: Andersen, Peter
; APPLICANT: Rosenkrands, Ida
; TITLE OF INVENTION: Therapeutic TB Vaccine
; FILE REFERENCE: SSI5AUSA
; CURRENT APPLICATION NUMBER: US/10/617,038
; PRIOR FILING DATE: 2003-07-11
; PRIOR APPLICATION NUMBER: DK PA 2002 01098
; PRIOR FILING DATE: 2002-07-13
; PRIOR APPLICATION NUMBER: US 60/401,725
; PRIOR FILING DATE: 2002-08-07
; NUMBER OF SEQ ID NOS: 187
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 28
; LENGTH: 393
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-617-038-28
```

```
Query Match 100.0%; Score 1990; DB 4; Length 393;
Best Local Similarity 100.0%; Pred. No. 2.1e-173;
Matches 393; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRDA1PLGR1AGFVNVNWSVTLVLTFTWSLATMLPGTVGGYPAVYVWLLGAGAWLL 60
   1 MRDA1PLGR1AGFVNVNWSVTLVLTFTWSLATMLPGTVGGYPAVYVWLLGAGAWLL 60
DB 1 MRDA1PLGR1AGFVNVNWSVTLVLTFTWSLATMLPGTVGGYPAVYVWLLGAGAWLL 60
   ASLAHELAHAHVARRAGVSVEYTLMLFGSVTLAGEAKTPKAAPRIAFAPATSLALS 120
QY 61 ASLAHELAHAHVARRAGVSVEYTLMLFGSVTLAGEAKTPKAAPRIAFAPATSLALS 120
   61 ASLAHELAHAHVARRAGVSVEYTLMLFGSVTLAGEAKTPKAAPRIAFAPATSLALS 120
DB 61 ASLAHELAHAHVARRAGVSVEYTLMLFGSVTLAGEAKTPKAAPRIAFAPATSLALS 120
   121 ATFGALAITTAGVTPPAIVISVAMWLTATVNLGLFNLLPGAPLDGRLVAAYLMRRHGD 180
QY 121 ATFGALAITTAGVTPPAIVISVAMWLTATVNLGLFNLLPGAPLDGRLVAAYLMRRHGD 180
   121 ATFGALAITTAGVTPPAIVISVAMWLTATVNLGLFNLLPGAPLDGRLVAAYLMRRHGD 180
DB 121 ATFGALAITTAGVTPPAIVISVAMWLTATVNLGLFNLLPGAPLDGRLVAAYLMRRHGD 180
   181 SVRGIGAAARGRVALVTLALGLAEFVAGLGVGWLAFIAGWTFPAAREEETRISTQQ 240
QY 181 SVRGIGAAARGRVALVTLALGLAEFVAGLGVGWLAFIAGWTFPAAREEETRISTQQ 240
   181 SVRGIGAAARGRVALVTLALGLAEFVAGLGVGWLAFIAGWTFPAAREEETRISTQQ 240
DB 181 SVRGIGAAARGRVALVTLALGLAEFVAGLGVGWLAFIAGWTFPAAREEETRISTQQ 240
   241 LFAAVRVADAMTAQPHTPAGVINEDEIQRVYVGERHSAAVPVADSDSITGLVALROLRD 300
QY 241 LFAAVRVADAMTAQPHTPAGVINEDEIQRVYVGERHSAAVPVADSDSITGLVALROLRD 300
   241 LFAAVRVADAMTAQPHTPAGVINEDEIQRVYVGERHSAAVPVADSDSITGLVALROLRD 300
DB 241 LFAAVRVADAMTAQPHTPAGVINEDEIQRVYVGERHSAAVPVADSDSITGLVALROLRD 300
   301 VAPRRSTSVGDIALPLHSVPTARPOEPLTALIERNAPLGPNSRALVTBGSAAVGVVTP 360
QY 301 VAPRRSTSVGDIALPLHSVPTARPOEPLTALIERNAPLGPNSRALVTBGSAAVGVVTP 360
   301 VAPRRSTSVGDIALPLHSVPTARPOEPLTALIERNAPLGPNSRALVTBGSAAVGVVTP 360
```

Db 301 VAPSRRTSTVGDIALLPHSVPTARPQEPFLTALLERMAPLGRSRALVTGSAVVGIVTP 360
Qy 361 SDVARLIDVYRLAOPETFTTSPQDADFPSDAG 393
Db 361 SDVARLIDVYRLAOPETFTTSPQDADFPSDAG 393

RESULT 2
US-10-047-260-12
; Sequence 12, Application US/10047260
; Publication No. US20020164706A1
; GENERAL INFORMATION:
; APPLICANT: Huang, Lisa
; APPLICANT: McCluskey, Michael
; APPLICANT: Larosaa, Robert
; TITLE OF INVENTION: HIGH LEVEL PROMOTERS FROM CYANOBACTERIA
; FILE REFERENCE: C11715 US NA
; CURRENT APPLICATION NUMBER: US/10/047,260
; CURRENT FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/264,925
; PRIOR FILING DATE: 2001-01-30
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: Microsoft Office 97
; SEQ ID NO: 12
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Synechocystis sp. strain PCC6803
US-10-047-260-12

Query Match 21.4%; Score 425.5; DB 4; Length 379;
Best Local Similarity 30.0%; Pred. No. 3.9e-30;
Matches 115; Conservative 70; Mismatches 147; Indels 51; Gaps 11;

Qy 1 MRDAIPLRIG--FVVVHVMSVLVILWFTWSLATML--PGTVGGYPAVVYMLLAGG 55
Db 11 MNNNRVSGSLRGIPFYVPSW--FLILGLVITLSYQDILARFPQLSGTR----WILGLIT 64
Qy 56 AVMLASLHAEHLAAVAVARRAGVSVEVTMLFGVGTALGGEAKTPKAAPRIAPAGPAT 115
Db 65 ALLPLASVVAHELGLSLVALAGIEVKSTLPLFGGLASLEKESWTQMAFVAIAGPAV 124
Qy 116 SLASATGALATITAGVTRPAIV--ISVAMWLAIVNLLGLFNLLPGAPLDGGRLVAY 173
Db 125 SLVL-----FLGLTIVGTQIPLPVGQAIIGLGMINIALFNLIPLGLDGGVLAISI 179
Qy 174 LMRHGDVSRAGIGARAGRVVALVIALGLAEFVAGLVGVMFLAFIGMFIFAAREEB 233
Db 180 VMQITGNONKGLILSRVGGGGMALAIISGLIINILPIGSEFTIILGMLIQNAGSSA 239
Qy 234 TRISTQQLFAGRVADAMTQOPHTAPGMINVEDFIQRYVLGER--HSAYPVADRDSITGL 292
Db 240 RNAQVKEQMEAFPTADAVIPNSPIIPAGINIREFANDVYIGKTPWRREFLVIGADNQLIGV 299
Qy 293 VALRLQRLVAVSRRTSTVGDIALLPHSVPTAR--POEPLTALLEMAPLGRSRALVTEG 351
Db 300 LATEDIK-----HVPISDWPQVTVDSIMQ-----PQMVTVAN 334
Qy 352 SAVVGIVTPSDVARLIDVYRLAQ 374
Db 335 QSLF-----EVAQLDDQKLSE 351

RESULT 3
US-10-156-761-14201
; Sequence 14201, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI

; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO: 14201
; LENGTH: 376
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-14201

Query Match 15.6%; Score 311; DB 4; Length 376;
Best Local Similarity 31.0%; Pred. No. 1.2e-19;
Matches 124; Conservative 47; Mismatches 173; Indels 56; Gaps 16;

Qy 7 LGRIAGFVVNV--HMSVL--VILWFTWSLATMLPGTVGGYPAVVYMLLAGAVMLLAS 62
Db 1 MGRPFQVPPVYVAPSMPLVAALITWVFGQLDRVLPEL-----GAARYLSLFFAVAFYAS 55
Qy 63 LIAHELAAVAVARRAGVSVEVTMLFGVGTALGGEAKTPKAAPRIAPAGPATSLASAT 122
Db 56 VLIHELHATVAALPFKLPVRRIQLQFPGVSEIEKETETPEREFVLAFFVGLSLIAGI 115
Qy 123 FGALATITAGVTRPAIVISVAMWLAIVNLLGLFNLLPGAPLDGRLVRAVLMRHRDSV 182
Db 116 FYLMAQIVDRKTVGVLLIA--GLMISNLVAAFNLLPGLPLDGGRLRAVVMYITGKPM 172
Qy 183 RAGIGAARAGRVVAL-VIALGL-----AEFVAG-----GLVGVMLAFIGW-- 223
Db 173 SGTIAAAMVGRALAVAVILGILPLTGTGALGSNAEDVGMDTVTDALLAI-LAAIITWG 231
Qy 224 ----FTPAAREESTRISTQQLFAGRVADAMTQOPHTAPGMINVEDFIQRYVGERHSA 279
Db 232 AGNSLRRARLEHLPRLRARLTR--RAVPVETDTPLS-----EALRRANDGAR--A 280
Qy 280 YPVADRDSITGLVALQLRDVAPSRRTSTVGDIALLPHS---VPTARPOEPLTALLER 336
Db 281 LVVADADGEPLSLREAAIVGPEHRKPWVDVSGLAODLTDGKMSLAGEDDLVDVRA 340
Qy 337 MAPLGRSRALVTEGSA--VVGIVTPSDVARLIDVYRLAQ 375
Db 341 T-----PATEYLVVESEGIYGVLSAADVERAF-VKAMARP 375

RESULT 4
US-10-156-761-14437
; Sequence 14437, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO: 14437
; LENGTH: 473
; TYPE: PRT

ORGANISM: Streptomyces avermectilis
US-10-156-761-14437

Query Match 7.2%; Score 143; DB 4; Length 473;
Best Local Similarity 27.7%; Pred. No. 0.0004;
Matches 98; Conservative 36; Mismatches 164; Indels 56; Gaps 17;

QY 67 ELAAHVAVRAGVSVE---VTLMFGVYALGGEAKTPKAFRIAFAPATSLASATF 123
DB 37 EKAAEGRRAGTVESELSKEISFQISG--TQGG---ITTSLVGMLAEPLADLLHGF 91
QY 124 GALATTAGTTPALVISVAMWLAT-VNLLGLFNLLP-----GAPLDGRLVRA--YIM 175
DB 92 TAIGPEGAV--PGAIVVGMGLASAVQVIG--ELVPMWAIKRPLOVAFVAPQHF 147
QY 176 RRGDSVRAGIGAAAGRVVALVIALG--AEFVAGLVGGVWALFTGWFIFAAAREE 233
DB 148 SRLFRPVIALMTV-ANRLVR---ALGVPEEEMASARTPSELVSLARHSAQALEDQ 202
QY 234 TRISTQDPAFGRVADAMTAQPHAP-----GWINVEDFIQRYVLGERHSAYPV-AD 284
DB 203 ----TADIFVRTLSIGELTAQHVTMPRVRSALQSSATAEDEV-NLUTRATGLSRFPVYRD 257
QY 285 RDGSGITGLVALRQLDVAAPSRRSTSVGDIALPLHSVTPARQEPULTALLERMAPLGPRS 344
DB 258 RIDFIVGMVHLKDALAIPSRLRTPVGRIAQPPILVPELTLFVQPLRLRSEQPIA--- 314
QY 345 RALVTEGSAVVGIVTPSDVARLI-----DVTYRLAQEPFTTSPDAD 387
DB 315 -VVDEYGGTAGVTLDEIVELVSGVRDEHDQDLPFLAVAPEDGRPAMDAD 367

RESULT 5

US-10-482-706-183
; Sequence 183, Application US/10482706
; Publication No. US20040241826A1
; GENERAL INFORMATION:
; APPLICANT: James, Brian William
; APPLICANT: Marsh, Philip
; APPLICANT: Hampshire, Tobias
; TITLE OF INVENTION: Mycobacterial Antigens Expressed During Latency
; FILE REFERENCE: 1581.1030000
; CURRENT APPLICATION NUMBER: US/10/482,706
; PRIOR FILING DATE: 2004-01-02
; PRIOR APPLICATION NUMBER: PCT/GB02/03052
; PRIOR FILING DATE: 2002-07-04
; PRIOR APPLICATION NUMBER: GB 0116385.6
; PRIOR FILING DATE: 2001-07-04
; PRIOR APPLICATION NUMBER: GB 0123993.8
; PRIOR FILING DATE: 2001-10-05
; NUMBER OF SEQ ID NOS: 282
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 183
; LENGTH: 591
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-482-706-183

Query Match 7.0%; Score 140; DB 5; Length 591;
Best Local Similarity 24.0%; Pred. No. 0.001;
Matches 111; Conservative 49; Mismatches 169; Indels 134; Gaps 23;

QY 10 IAGFVNVNH---WS---VLVILMLF-----TMSLATMLDGTG---GVP 44
DB 70 VAGAIIRVAPLWTPAPISLVVQLLESIALRALYIVSSMRPULLITPLPPLAVP 129
QY 45 AVVYIILGAGAVMLASILAHELAHAVVARAGVSVEVTLMFGVYALGGEAKTPKA 104
DB 130 GFAMWAAALNSLPLMAA--LAWVCADAILVLTGHHRAVLT---GVLVYIGGLFFPEKA 183
QY 105 AF--RIARAGPATSLALSTFGALA-ITLACVR--TPRAIVSVAN---MLATV----- 149
DB 184 AVIPVSPAVVALQCHVRGDRSALATVWRAGVRLMTPSLATLVGVVALYLAVVDQRWSS 243

QY 150 -----NLILGLEFNLLPGAPLDGRLVRAVYLMRRHDSVRAGIGAAPAGRVVA-- 196
DB 244 DLSMTWDLCSVTHIGVIALAGGFWMDWARWAPSPW-----ATPPAVVWVL 290
QY 197 --LVILNAGLAEFAAGLVGGVWLAFTGWFIFAAAREETRISTQQLFAGRVADAMTAQ 254
DB 291 GMLVILIALVALSLVRKRRIIGVWMLTAAG--YAAVACQPIFLMSSPFTALELAQTLYF 347
QY 255 PHTAGWINVEDFIQRYVLGERHSAYPVADRDGSTITGVALLRQLDVAAPSRRSTSVGDI 314
DB 348 P-----DLV--VVL-----ALLAAVALQ-----APNRAGRMLD-- 374
QY 315 ALPLHSVPT-ARPEPITALLERMAPLGRSARLVTEG---SAVGIYTPSDVARLID-- 368
DB 375 ASPAAVAVTASAVLFTLSLYSTATFLASWRDNPTEGYLKNQAQSLAAASGAPLLDQE 434
QY 369 -----VYRLAOPE-----PFTTSPDADRFSDAG 393
DB 435 VDPVLQRAVAMPENLASHMFLALRLVRPFAVTTQLRMFTSTG 477

RESULT 6

US-10-080-170-24
; Sequence 24, Application US/10080170
; Publication No. US20030123601A1
; GENERAL INFORMATION:
; APPLICANT: COLE, S. T.
; TITLE OF INVENTION: COMPARATIVE MYCOBACTERIAL GENOMICS AS A TOOL FOR
; IDENTIFYING TARGETS FOR THE DIAGNOSIS, PROPHYLAXIS OR
; TREATMENT OF MYCOBACTERIOSES
; FILE REFERENCE: 03495.0218
; CURRENT APPLICATION NUMBER: US/10/080,170
; PRIOR FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 60/270,123
; PRIOR FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 652
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 24
; LENGTH: 649
; TYPE: PRT
; ORGANISM: Mycobacterium leprae
US-10-080-170-24

Query Match 7.0%; Score 140; DB 4; Length 649;
Best Local Similarity 23.5%; Pred. No. 0.0011;
Matches 114; Conservative 62; Mismatches 172; Indels 138; Gaps 25;

QY 11 AGFVNVNHSVYLILM--LFTWSL---ATMLP-----GTVG 41
DB 105 AALVLSVLGVLMLGTGRLVAPSLQGRQAIMLPAGALVYVALPPARDFATSGLESGLVL 164
QY 42 GYPAVVTYL-----LGAGAVMLASILAHELAHAVVARAGVSVEV 84
DB 165 TYLGILMWMWVCMAPLNRQSRRFIGALAFVAGCSVLVPELALMGSSALIMLMTAAR 224
QY 85 TLMFGVYALGGEAKTPKAFRIAFAG---PATSLASAT---FGALATITLACVTRPAI 138
DB 225 TCMIRALIVAGGSLPAVAYQLFRKGYGLVPGTALAKDAGDKMSQGITVLSNPNQPV 284
QY 139 VISVAMWLATVNL-LGLFNLL-----PGAPLDGR-----L 169
DB 285 L-----WVPLVILVLGLLMLIRHPSFMHPLFTPDGGRVARAVQSPPAVVFVFSGL 339
QY 170 VRATLMRR-----HGDSVRAGIGAAAGRVVALVIALGLAEF--VAGSLVGV---W 217
DB 340 LQAFYMWIRQGGDFMHRGVTLAPLCLLAPVVIVIVISEG--ADFSRQTNWLAGVTSILW 398
QY 218 LAFIGWFIFAA---AREETRISTQQLFAGRVADAMTAQ--HTAPGWINVEDFIQRY 271
DB 399 LGVAGKSLMAANSQMGDDATNVS-----YSGIVDERFRYAQATGHAP--LTAADYL--- 449
QY 272 VLGERHSAYPVADRDGSTITGVALLRQLDVAAPSRRSTSVGDIALPLHSVTPARQEPBLT 331

```
Db 450 -----GYP---RMAAV--LVALLNTPDGAALLPSSGYIKMDLVPMTQLSPSSGSPD 497
Qy 332 ALBERMAPLGRSRALVTESGAVVIGTVPDVARLIDVYRLAQPEPTTSPQ-----D 385
Db 498 SLVSQK----PQHTVFET-NLGMIGMNVGLDV-RVIDQIGLANPLAQHTERLQHRIGHD 551
Qy 386 ADRESD 391
Db 552 KNLFPD 557

RESULT 7
US-10-080-170-24
; Sequence 24, Application US/10080170
; Publication No. US20040121322A9
; GENERAL INFORMATION:
; APPLICANT: COLE, S.T.
; TITLE OF INVENTION: COMPARATIVE MYCOBACTERIAL GENOMICS AS A TOOL FOR
; IDENTIFYING TARGETS FOR THE DIAGNOSIS, PROPHYLAXIS OR
; TREATMENT OF MYCOBACTERIOSES
; FILE REFERENCE: 03495, 0218
; CURRENT APPLICATION NUMBER: US/10/080, 170
; PRIOR FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 60/270, 123
; PRIOR FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 652
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 24
; LENGTH: 649
; TYPE: PRT
; ORGANISM: Mycobacterium leprae
US-10-080-170-24

Query Match 7.0%; Score 140; DB 4; Length 649;
Best Local Similarity 23.5%; Pred. No. 0.0011;
Matches 114; Conservative 62; Mismatches 172; Indels 138; Gaps 25;

Qy 11 AGPVNNVMSVYLVLIM--LFTWSL---ATMLP-----GTVG 41
Db 105 AALVLSVIGWVLMGLTGRLVAPSLQGRQAIMLPAGALVYVALPPAPDFATSGLESGLV 164
Qy 42 GYPAVVVWL-----LGAGAVMLASLALHELAAVAVARRAGVSVESV 84
Db 165 TYLGILMMWVWCMAQPLNRSSQSRFFIGALAFVAGCSVLVRPELALMGSGALIMIMIAAR 224
Qy 85 TLMFGVTALGGEKTPKKAFLIAFAG---PATSLASAT---FGALAITLAGVTRPAI 138
Db 225 TCWLRALIVVAGSGLPVAYQLFRMGVYGLVPGTALAKDAAGDKWSGCIITLISFNQPYV 284
Qy 139 VISVAMMLATVNL--LGLFNLIL-----FGAPLDGGR-----L 169
Db 285 L-----WPLVLLVLGLLMLLHHRWSPFMHPLETPDGRVAVAVOSPFAVVVFVPSGL 339
Qy 170 VRAYLMRR-----HGDSVRAGIGAAAGRVVALVLIAGLAEF--VAGLVGV---W 217
Db 340 LQAFWTLRQGGDFMHRGVLALPLFCLLAPVVVLPVVISG--ADFSRQTGMWLAGVTSLLM 398
Qy 218 LAFIGWFIFAA---AREEETRISTQQLFAGVRVADAMTAQP--HTAPGMINVEDFIORY 271
Db 399 LGVAGWMLMAANSRGMGDDATNVS---YSGIVDERRFYAQATGHARP--LTAADYL--- 449
Qy 272 VLGERHSAVYVADBDGSIIGLVALRQLRDVAPSRRSTTSVGDIALPLHSVPTAPPOEPLT 331
Db 450 -----GYP---RMAAV--LVALLNTPDGAALLPSSGYIKMDLVPMTQLSPSSGSPD 497
Qy 332 ALBERMAPLGRSRALVTESGAVVIGTVPDVARLIDVYRLAQPEPTTSPQ-----D 385
Db 498 SLVSQK----PQHTVFET-NLGMIGMNVGLDV-RVIDQIGLANPLAQHTERLQHRIGHD 551
Qy 386 ADRESD 391
Db 552 KNLFPD 557
```

```
Db 450 -----GYP---RMAAV--LVALLNTPDGAALLPSSGYIKMDLVPMTQLSPSSGSPD 497
Qy 332 ALBERMAPLGRSRALVTESGAVVIGTVPDVARLIDVYRLAQPEPTTSPQ-----D 385
Db 498 SLVSQK----PQHTVFET-NLGMIGMNVGLDV-RVIDQIGLANPLAQHTERLQHRIGHD 551
Qy 386 ADRESD 391
Db 552 KNLFPD 557

RESULT 8
US-10-468-356-24
; Sequence 24, Application US/10468356
; Publication No. US20040197896A1
; GENERAL INFORMATION:
; APPLICANT: COLE, STEWART
; TITLE OF INVENTION: COMPARATIVE MYCOBACTERIAL GENOMICS AS A TOOL FOR
; IDENTIFYING TARGETS FOR THE DIAGNOSIS, PROPHYLAXIS OR
; TREATMENT OF MYCOBACTERIOSES
; FILE REFERENCE: 05394, 0019
; CURRENT APPLICATION NUMBER: US/10/468, 356
; PRIOR FILING DATE: 2003-08-19
; PRIOR APPLICATION NUMBER: 10/080, 170
; PRIOR FILING DATE: 2002-02-22
; PRIOR APPLICATION NUMBER: 60/270, 123
; PRIOR FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 655
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 24
; LENGTH: 649
; TYPE: PRT
; ORGANISM: Mycobacterium leprae
US-10-468-356-24

Query Match 7.0%; Score 140; DB 4; Length 649;
Best Local Similarity 23.5%; Pred. No. 0.0011;
Matches 114; Conservative 62; Mismatches 172; Indels 138; Gaps 25;

Qy 11 AGPVNNVMSVYLVLIM--LFTWSL---ATMLP-----GTVG 41
Db 105 AALVLSVIGWVLMGLTGRLVAPSLQGRQAIMLPAGALVYVALPPAPDFATSGLESGLV 164
Qy 42 GYPAVVVWL-----LGAGAVMLASLALHELAAVAVARRAGVSVESV 84
Db 165 TYLGILMMWVWCMAQPLNRSSQSRFFIGALAFVAGCSVLVRPELALMGSGALIMIMIAAR 224
Qy 85 TLMFGVTALGGEKTPKKAFLIAFAG---PATSLASAT---FGALAITLAGVTRPAI 138
Db 225 TCWLRALIVVAGSGLPVAYQLFRMGVYGLVPGTALAKDAAGDKWSGCIITLISFNQPYV 284
Qy 139 VISVAMMLATVNL--LGLFNLIL-----FGAPLDGGR-----L 169
Db 285 L-----WPLVLLVLGLLMLLHHRWSPFMHPLETPDGRVAVAVOSPFAVVVFVPSGL 339
Qy 170 VRAYLMRR-----HGDSVRAGIGAAAGRVVALVLIAGLAEF--VAGLVGV---W 217
Db 340 LQAFWTLRQGGDFMHRGVLALPLFCLLAPVVVLPVVISG--ADFSRQTGMWLAGVTSLLM 398
Qy 218 LAFIGWFIFAA---AREEETRISTQQLFAGVRVADAMTAQP--HTAPGMINVEDFIORY 271
Db 399 LGVAGWMLMAANSRGMGDDATNVS---YSGIVDERRFYAQATGHARP--LTAADYL--- 449
Qy 272 VLGERHSAVYVADBDGSIIGLVALRQLRDVAPSRRSTTSVGDIALPLHSVPTAPPOEPLT 331
Db 450 -----GYP---RMAAV--LVALLNTPDGAALLPSSGYIKMDLVPMTQLSPSSGSPD 497
Qy 332 ALBERMAPLGRSRALVTESGAVVIGTVPDVARLIDVYRLAQPEPTTSPQ-----D 385
Db 498 SLVSQK----PQHTVFET-NLGMIGMNVGLDV-RVIDQIGLANPLAQHTERLQHRIGHD 551
Qy 386 ADRESD 391
Db 552 KNLFPD 557

RESULT 9
US-10-282-122A-50265
; Sequence 50265, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
```

```

; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zykend, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Foreyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034a
; CURRENT APPLICATION NUMBER: US/10/282,122A
; PRIOR FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 50265
; LENGTH: 706
; TYPE: PRT
; ORGANISM: Burkholderia mallei
US-10-282-122A-50265

Query Match      7.0%; Score 138.5; DB 4; Length 706;
Best Local Similarity 22.2%; Pred. No. 0.0018;
Matches 110; Conservative 48; Mismatches 159; Indels 179; Gaps 21;

QY 7 LGRIAGFVVNV---HMSVLVILMLFTWSLATMLPGTVGGYPAVVVYMLGAGGAVMLLASL 63
DB 121 LGLIAMLVPLALAGHSAV-PLIFVCGIAAACVSAAGGWRATPLRLIAGSVCMILFSA 179
QY 64 LAHEI-----AAHVARRAGVSVEVTLMIFGGVTALG-----GE 98
DB 180 VTTLLAEFEQTVVGASIMASGLYOPGAAGRDALLMLVAPLALPLVIRPLDPLALGD 239
QY 99 AKTPKAAIRIAFAGPATSLA---LSATRGALAITTAGVTRPAIVISVAMWMLATVLL--- 152
DB 240 DAAAAAGRV---DATRLAGTVAVVGFAVSAVSIAG---PLSYIG---LIAPNLIRQM 288
QY 153 -----LGLFNLLPGAPLDGRLV----- 170
DB 289 RGAKASRIRGA--LVFLAALAGALVLVVDSANVLALGLDALTSTGVAIAFVGTRPLMLAMIR 346
QY 171 RAYLW---RRHGDVPRAGI--GAARAGR-----VVALVTLALGLAEFVAGLVGVWL 218
DB 347 RGAASGAAGQOPAHYRAGTRGAGRLVRAIAVLVPMPAVAVLLLAAGAAVLVVGASFGPMWI 406
QY 219 AFIGFIFAARBEETR-----ISTQOLF----- 242
DB 407 GGRFFAALAQDDIDARVTLDIRAPRLCALLAGALLGASGVLMQSIIVRNPLAGEBVLGV 466
QY 243 ---AGVRVADMTAQPHTAQWINVED-----FIQRYVLGERHSAYPVADRDGSITGL 292

```

```

DB 467 TQAGIATVATVAMPPLAAHGLVLAASLAGGATVLTALNRRRVAPLA---VALTGI 523
QY 293 VA-----LQRLRVAPSRSTTSVGD-----IALPIHVPTR 325
DB 524 VIGTILWTLAOWLITQESVQPARFVVLVGGTGYRSGEAAALLPFGVLAVALAL-LAR 582
QY 326 POEPLTALLERMAPLG 341
DB 583 PDLIALGDDEQAALAG 598

RESULT 10
US-10-156-761-9868
; Sequence 9868, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 9868
; LENGTH: 570
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-9868

Query Match      6.7%; Score 133.5; DB 4; Length 570;
Best Local Similarity 24.5%; Pred. No. 0.0038;
Matches 105; Conservative 46; Mismatches 163; Indels 115; Gaps 22;

QY 10 IAGFVVNVHMSVLVILMLFTWSLATMLPGTVGGY-----PAVVYV-LIGA 53
DB 190 LAGTV--VLWMLPLPRRSATVLCGRGLVPSIARYGPVAVVALALLIPVVRPRMG 247
QY 54 GG-----AVMLLASLAHELAHVAVARAGVSVEVTLMIFGGVTALGGEAKTPKAAE- 106
DB 248 GGMTDWRRACTLGAACAAQQLPW-----FLTPFLAGVYALRGERGPRALVV 295
QY 107 ---RIAPGPTSLSLSTFGALAITTAGVTRPAIVISVAMWMLATVLLGLFNLLPGAP 163
DB 296 AVGRIGAVAAAGTQWLLINAF-----VSEPRSMLSGI-----MLPLTQSAV 336
QY 164 LDGGLVAYVIMRHBGSVVRAGIQAARAGRVVALVTLALGLAEFVAGLVGV--WLAFI 221
DB 337 IHGGLVGASLYTYDSSGRLLAW--SHASTILAAVLALVLFMRRLGPRATVLPKCF- 393
QY 222 GWFIFAAREBEETRIISTQOLFAGVRVADMTAQPHT-APGMINVEDFIQRYVLGERHS- 278
DB 394 ----FLATRSQD---GYLLMTPLMLAAAVTAPSAFASAW-----QPRLPFGAHRRA 440
QY 279 -----AYPADRDGSLTIGVALRQLRDVAPSRSTTSVGDIALPLHSV--PRARP 326
DB 441 ARVGVALLLALLPLSAAYAVTGAAPPLRM--EVALAVRRGTGTGTLRLRVNTGTTMRP 498
QY 327 QEPLTAL--LER--MAPLGRSRALVTEGSAVVGIVTSPDVARLIDVYRLAQPEP----- 377
DB 499 HFTVTTMOGMRKYSVAGAPAS--VPAHGSARYELPLPDG-----RFLPLPFGVIR 548
QY 378 --FTTSPO 384

```

Db 549 LRAFSASPE 557

RESULT 11
US-10-493-462-25
; Sequence 25, Application US/10493462

ORGANISM: Mycobacterium tuberculosis
US-10-493-462-25

```

; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 14
; LENGTH: 3745
; TYPE: prt
; ORGANISM: micromonospora carb
US-10-205-032-14

```

Query Match	6.6%;	Score 132;	DB 4;	Length 3745;
Best Local Similarity	25.2%;	Pred. No. 0.066;		
Matches 107;	Conservative 40;	Mismatches 174;	Indels 104;	Gaps 20;

RESULT 12
US-10-205-032-14
Sequence 14, Application US/10205032

ORGANISM: Streptomyces halstedii sp. HC-34
US-10-378-083-6

Query Match	6.6%;	Score 130.5;	DB 4;	Length 3362;
Best Local Similarity	24.1%;	Pred. No. 0.079;		
Matches 87;	Conservative 52;	Mismatches 153;	Indels 69;	Gaps 14;

```

Db      471 IAGSGVTPPWVLSABSADALRCQABERLISFVSAAQDVSVDDVAYSLSGVSRAGLERHGVV 530
Qy      141 SVAMWMLATVNLILGTFNLLPGAPLDC-----GRLVAYLMRRHGDVRAIGIAARAG 192
Db      531 GE-----SRRELLAALESLSAGVSRGVTVGRAEGRU--AFLPTGOG-AQRGMKRELA- 582
Qy      193 RVVALVLIALGLAEFVAGLVGVWMLAFIWFIFPAAREBEETRISTOOLPAGVRVADAMT 252
Db      583 --AEPFLFAASLE--TCGLLERAGVA--VREVLFAEBSGSAEAAALLTRVYAOALFAVEV 637
Qy      253 AQPHTAPGIVNEDFIQRVVGERHSAY--PVADDGSLTGVALRQLRDVAPSRKSTIS 310
Db      638 ALFRLVESFGVVPDFVAGHSGVEIAAAHYAGVSLSDAVSLVAABGRMLDALPEGAMVA 697
Qy      311 VGDIALPLHSVPTAPRQEPFLTALLERMAPLGRSRALVTESGAVVGIPTSPDVARLIDVY 370
Db      698 V-----QATEEDVLLALLEGVE---DASIAIANGPDAVVVSGTEAGVAAVVDVL 742
Qy      371 R 371
Db      743 R 743

```

```

RESULT 14
US-10-156-761-10433
; Sequence 10433, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 10433
; LENGTH: 4685
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-10433

```

```

Query Match      6.6%; Score 130.5; DB 4; Length 4685;
Best Local Similarity 23.6%; Pred. No. 0.12;
Matches 118; Conservative 47; Mismatches 168; Indels 167; Gaps 26;
Qy      7 LGRIAGFVNVVMSVLTVMILFTMSLATVLP-----TVGGYPAVVVM----- 49
Db      997 VGRLSG-----TWLVVT--PGGAPVEEVRALAVGAEAIVLWKEBE 1038
Qy      50 -----LLGAGAVMLLASILAHLELAHVVA-----RRAGVESVETLM-LFGVTA 94
Db      1039 GREALAGRLTGA-GGVSGVSLIGWDEEAAPVATVTLVQALDADGVEA-PLVVLTOGAAS 1096
Qy      95 LGGEAKTRKAARIPAPGATSLASATPGALATILAGVTRTPATYISVAMWMLATVNLILG 154
Db      1097 VGAEEVVPHPVOTQVVALQVAGLEQPGSGVGL-VDPGVWDERVASGLTAVLAA----- 1149
Qy      155 LFNLLPGAPLD-----GRLVRAVYL-----WRRHGSV-----RAGIGA--- 188
Db      1150 -----GEGEDOVAVRSSGAYARRLVRAPLGANPAPVADNMHGVTLITGGGGIGAHIA 1203
Qy      189 ---ARAGRVVALVLI-----ALGLAEFVA--GGLVGVWMLAFIGWFIPAAREBEETRIS 237

```

```

Db      1204 RMLAKEG-AERLLVSRSGEQAEAGAEIATELGL--GAETVF-----AACDVSRDA 1253
Qy      228 TOOLFAGVRVADAMTAQPHTA--PGMINV-----EDFIQRYVLGERH----- 277
Db      1254 LAHVYAGIPAEHPLTAVHTTGISYAEIATATPEHFDVLSARVLGARHIDELTAELEGV 1313
Qy      278 --SAYFV-----ADRGSLTGLVALROLRDVADS-----RSTSVG 312
Db      1314 ELEAFVVFSSGAAVWVGSSNGANMAAGYLDGLVRRRARAGVAGTSVSGWQATAMAVG 1373
Qy      313 DIALPLHSVPTAPRQEPFLTALLERMAPLGRSRALVTESGAVVGIPTSPDVARLIDVRYL 372
Db      1374 DTAEGU--SRRGVRLIDPVLAIV-----QALROVLEQDEVSVTVTMDMALFTPGYAM 1423
Qy      373 AQPHTFTTSPQDADRFSDA 392
Db      1424 ARRRRLIEDIPEAGALSQA 1443

```

```

RESULT 15
US-10-437-963-177542
; Sequence 177542, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; PRIOR FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 177542
; LENGTH: 487
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_75183C.1.pep
US-10-437-963-177542

```

```

Query Match      6.4%; Score 128; DB 4; Length 487;
Best Local Similarity 22.7%; Pred. No. 0.0097;
Matches 54; Conservative 33; Mismatches 93; Indels 58; Gaps 9;
Qy      52 GAGGAVMLLASILAHLELAHVAVRAGVS-----VESVTLMFGGVTAALGGEAKTPKAA 105
Db      241 GVGAGLVAAITIGVHEIHAIILAADTGIKLAVPYFVPSWQISFGAIRTIVNIYANREDL 300
Qy      106 FRIAPGATSLAV-----SATFGALATTLG-----V 133
Db      301 LKVAAGAPLAFSGFVLLILGFTLPSPDGLVVIDPAVPHESFLVGLATLILGDALKE 360
Qy      134 RTPAIVISVAMWMLATVNLILGTFNLLPGAPLDGRLVRAVLMRRHGDVRAIGIAARAGR 193
Db      361 GTKLSINPLVLM-AWAGILLVAINSIPAGEIDGGRIAFV-AW-----GRKISSR 407
Qy      194 VVALVLIALGLAEFVAGLVGVWMLAFIGWFI-----PAAAREBEETRISTOOLPAGVRY 247
Db      408 ISSIATIGLISALFND--VAFYVVL--FPLQKGPISPLSEETPEPNNYISIGVAI 462

```

Search completed: March 23, 2006, 06:28:17
Job time : 84.6454 secs

This Page Blank (usp)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: March 23, 2006, 06:17:23 ; Search time 9.43101 Seconds
(without alignments)
1192.746 Million cell updates/sec

Title: US-10-617-038-28

Perfect score: 1990

Sequence: 1 MRDAIPGRINGFVVNVHMS.....QPEPTTSPQDADRFSDAG 393

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 169630 seqs, 28622889 residues

Total number of hits satisfying chosen parameters: 169630

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA.New.*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US03_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	121	6.1	224	7	US-11-098-686-10266
2	119	6.0	529	6	US-10-858-730-104
3	119	6.0	529	6	US-10-858-730-105
4	118.5	6.0	8695	7	US-11-205-109-15
5	115	5.8	403	7	US-11-205-109-23
6	112.5	5.7	3655	7	US-11-075-185-5
7	112	5.6	511	7	US-11-087-099-11617
8	110.5	5.6	497	7	US-11-143-980-61
9	109	5.5	488	6	US-10-467-657-2362
10	109	5.5	514	7	US-11-087-099-1493
11	109	5.5	619	7	US-11-205-109-30
12	106.5	5.4	334	6	US-10-858-730-121
13	105.5	5.3	434	7	US-11-087-099-6812
14	104.5	5.3	488	7	US-11-087-099-3033
15	104.5	5.3	557	7	US-11-087-099-9168
16	104	5.2	424	7	US-11-087-099-5604
17	104	5.2	3689	7	US-11-075-185-4
18	103.5	5.2	212	7	US-11-212-443-48
19	102.5	5.2	477	7	US-11-087-099-2487
20	102	5.1	377	7	US-11-087-099-12099
21	102	5.1	395	7	US-11-087-099-11975
22	102	5.1	5712	7	US-11-143-980-47
23	102	5.1	7968	7	US-11-143-980-49
24	101.5	5.1	423	7	US-11-096-568A-25217
25	101	5.1	324	6	US-10-858-730-113

26	101	5.1	617	7	US-11-143-980-35	Sequence 35, Appl
27	100.5	5.1	232	6	US-10-467-657-3352	Sequence 3352, Ap
28	100.5	5.1	563	7	US-11-087-099-6896	Sequence 6896, Ap
29	100	5.0	542	6	US-10-858-730-106	Sequence 106, App
30	99	5.0	471	7	US-11-087-099-4957	Sequence 4957, Ap
31	99	5.0	476	7	US-11-087-099-12104	Sequence 12104, A
32	99	5.0	483	7	US-11-087-099-3687	Sequence 3687, Ap
33	99	5.0	573	7	US-11-087-099-7597	Sequence 7597, Ap
34	98.5	4.9	465	7	US-11-087-099-8665	Sequence 8665, Ap
35	98.5	4.9	507	6	US-10-479-873-7	Sequence 7, Appl
36	98.5	4.9	507	6	US-10-479-873-8	Sequence 8, Appl
37	98.5	4.9	578	6	US-10-858-730-103	Sequence 103, App
38	98.5	4.9	595	7	US-11-087-099-3825	Sequence 3825, Ap
39	98.5	4.9	613	6	US-10-467-657-5796	Sequence 5796, Ap
40	98	4.9	415	7	US-11-205-109-20	Sequence 20, Appl
41	98	4.9	429	7	US-11-205-109-32	Sequence 32, Appl
42	98	4.9	511	7	US-11-087-099-8192	Sequence 8192, Ap
43	97.5	4.9	574	7	US-11-261-228-2	Sequence 2, Appl
44	97	4.9	324	6	US-10-858-730-120	Sequence 120, App
45	97	4.9	478	7	US-11-087-099-1476	Sequence 1476, Ap

ALIGNMENTS

```
RESULT 1
US-11-098-686-10266
; Sequence 10266, Application US/11098686
; Publication No. US20060024696A1
; GENERAL INFORMATION:
; APPLICANT: Kapur, Vivek and Gebhart, Connie J.
; TITLE OF INVENTION: NUCLEIC ACID AND POLYPEPTIDE SEQUENCES
; FILE REFERENCE: 09531-128001
; CURRENT APPLICATION NUMBER: US/11/098,686
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: PCT/US03/31318
; PRIOR FILING DATE: 2003-10-01
; PRIOR APPLICATION NUMBER: US 60/416,395
; NUMBER OF SEQ ID NOS: 11433
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10266
; LENGTH: 224
; TYPE: PRT
; ORGANISM: Lawsonia intracellularis
US-11-098-686-10266

Query Match      6.1% Score 121; DB 7; Length 224;
Best Local Similarity 27.0%; Pred. No. 0.025;
Matches 54; Conservative 21; Mismatches 75; Indels 50; Gaps 7;

QY 57 VMLASLAHLELAHVAARRAGVSVEVTLMDFGVT-----ALG----- 96
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 18 VPLVIGIILHEHVAHGWAVQRRG---DPTAAMLGITINPTHHIDPLGMVFTITSLSP 73
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 97 ---GEAK-----TPKAAPRIAPAGPATSLASATFG-ALATITLAVGRTPATVYS 141
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 74 FIFGAKAVVPINPRFYLRKDTMTLVSPAGSRAANFLAFTGLRLITLFFPAVEMQOS 133
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 142 VAWML-----ATVNLLIGLFINLLPGAPLDGGLVAVYIMRRHGSVRAIGARAGR 193
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 134 SVWTFYFLMFSSGIIINFTLAWMLNLPPIPIIDSGSKITWALIPNELGYKML---AERYGF 190
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 194 VVALVTLLALGLAEFVAGSLV 213
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 191 VIFALLLTGTLGTYLPLI 210
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

RESULT 2
US-10-858-730-104
; Sequence 104, Application US/10858730
; Publication No. US20050255568A1
```

```

; GENERAL INFORMATION:
; APPLICANT: Bailey, Richard B.
; APPLICANT: Blomquist, Paul
; APPLICANT: Doren, Reed
; APPLICANT: Driggers, Edward M.
; APPLICANT: Madden, Kevin T.
; APPLICANT: O'Leary, Jessica
; APPLICANT: O'Toole, George
; APPLICANT: Trueheart, Joshua
; APPLICANT: Walbridge, Michael J.
; APPLICANT: Yorgey, Peter S.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR AMINO ACID
; TITLE OF INVENTION: PRODUCTION
; FILE REFERENCE: 14184-030001
; CURRENT APPLICATION NUMBER: US/10/858,730
; CURRENT FILING DATE: 2004-06-01
; PRIOR APPLICATION NUMBER: US 60/475,000
; PRIOR FILING DATE: 2003-05-30
; PRIOR APPLICATION NUMBER: US 60/551,860
; PRIOR FILING DATE: 2004-03-10
; NUMBER OF SEQ ID NOS: 364
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 104
; LENGTH: 529
; TYPE: PRF
; ORGANISM: Mycobacterium tuberculosis
; JS-10-858-730-104

```

Query Match	6.0%;	Score 119;	DB 6;	Length 529;
Best Local Similarity	27.1%;	Pred. No. 0.098;		
Matches 105;	Conservative 32;	Mismatches 143;	Indels 108;	Gaps 21

Qy	1	AGGVNVAHMSVLYILFLPFWLSATMLPGVG-----GYPAVVYWMLGAGGA--VML	59
Db	166	AGFALGVA--MLGGTWL--TCVLAANTSGVIDRLGRLKRISTEPLPQRFVGAGIATIVAV	223
Qy	60	LASLHAHELAHVAVARAGVSVEVTMLFGGVTALGG--EAKTPKAFFIAFGAPATSL	117
Db	224	AAVLINGDPTALVA--TGIVV-----LLSGMTLVGMDQAVTGVMLTALARIADALFL	275
Qy	118	ALSATGALAITTAGYRTPAIVISV---AMWLATVNLGLFNLLFGALPLDGRLL--VRA	172
Db	276	TAGIVVGIL--ISLRGTNAGIQIEIENVDATTTLAPGPEPLVAVSGAALLSGVCLTIAS	334
Qy	173	YLMRRHGDSVRAGIGABAGRVVALVTILALAEFVAGALGVWMLFICGFIIPAAREE	232
Db	335	YAPLR--SVATKGLSAGLA---ELVLLGLGAAGF--GRVVAITWTALIGVFIA-----	380
Qy	233	ETRISTQOLFAGRVVADAM--TAQPTAPRWINVEDFIQRYVLGERHSAYPAVDRDSITG	291
Db	381	-TISLR-----RQAPALVTATAGIMP-----MLPG	405
Qy	292	LVALROLRDVAPSRNSTSVGDIALPLHSVTPARQEPYTLALLERMAPLGRSRALVTBG	351
Db	406	LAVFRAVFAFA-----VND-----TPDGGTLTLLAA-----TALLAGS	440
Qy	352	SAVVGIVTPSDV---ARLIDVYRLAQP	375
Db	441	GVVLAGEFLASPLYYGAGRIQGLDFRLEGF	468

RESULT 3
US-10-858-730-105
; Sequence 105, Application US/10858730
; Publication No. US20050255568A1

1 APPLICANT: Bailey, Richard B.
 2 APPLICANT: Blomquist, Paul
 3 APPLICANT: doten, Reed
 4 APPLICANT: Driggers, Edward M
 5 APPLICANT: Midden, Kevin T.
 6 APPLICANT: O'Leary, Jessica
 7 APPLICANT: O'Toole, George

```

/ APPLICANT: Trueheart, Joshua
/ APPLICANT: Walbridge, Michael J.
/ APPLICANT: Yorgey, Peter S.
/ TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR AMINO ACID
/ TITLE OF INVENTION: PRODUCTION
/ FILE REFERENCE: 14184-030001
/ CURRENT APPLICATION NUMBER: US/10/858,730
/ CURRENT FILING DATE: 2004-06-01
/ PRIOR APPLICATION NUMBER: US 60/475,000
/ PRIOR FILING DATE: 2003-05-30
/ PRIOR APPLICATION NUMBER: US 60/551,860
/ PRIOR FILING DATE: 2004-03-10
/ NUMBER OF SEQ ID NOS: 364
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 105
/ LENGTH: 529
/ TYPE: prt
/ ORGANISM: Mycobacterium tuberculosis
/ US-10-858-730-105

```

Query Match	6.0%;	Score 119;	DB 6;	Length 529;
Best Local Similarity	27.1%;	Pred. No. 0.098;		
Matches 105;	Conservative 32;	Mismatches 143;	Indels 108;	Gaps 21;

```

OY      1  AGGVVAVVHNSVLTIVLMLFTWLSLATMLPGVNG-----GYPAVYVYMLGAGGA--VML 59
Db      166  AGGALGVA--MLGGTVM--TCVLAATSGVIDRLGRLLNRIGTFLPFRVFGAGATLTIVAV 2233
OY      60  LASLSLAHELAHAAVVARAGVSVESTVLMFGGVTALGG--EAKPKPAAFRIAEPATSL 117
Db      224  AAVLIAGDPTALVA--TGIVV-----LTSMTLVGSMQDAVYGVWLTALARGDALFL 275
OY      118  ALSATGALATITAGVTPAIVISV---AMWLATVNLGLFNLLPGAPLDGRL--VRA 172
Db      276  TAGIIVGIL--ISLRGTNAGIOIEIHDVATTTLAPGMPFLILVAVGAALSGVCLTIAS 334
OY      173  YLMRRHGDSDVRAGIGABAGRVVALVITLALGLAEFVAGAGVGWMLPFGMFLPAAREE 232
Db      335  YAPLR--SVATNGLSAGLA---ELVILGLGAGF--GRVVAVTMALIGVGFLA----- 360
OY      233  ETRISTCOLFAGRVVDAM--TAQPHATAGMINVEDFIORVVLGERSASYPVARDGDSITG 291
Db      381  -TLISIR-----RQAPALVTATAIMP-----MLPG 405
OY      292  LVALRQLRDVAPSRSTTSVGDIALPLHSVPTARQEPVLTALLERMAPLGRSALVTBG 351
Db      406  LAFFRAVFAFA-----VND-----TFDGGITQLLEAAA-----TALLAGS 440
OY      352  SAVVGIVTPSDV-----ARLIDVYRLAQ 375
Db      441  GVVLAGFELASLPYGGAGRIGDILFRIEGP 466

```

RESULT 4
US-11-205-109-15

Sequence 15, Applicant US 1125305
Publication No. US20050287641A1
GENERAL INFORMATION:
APPLICANT: Farnet, Chris
APPLICANT: Zazopoulos, Emmanuel
APPLICANT: Staife, Alfredo
TITLE OF INVENTION: GENE CLUSTER FOR RAMOPLANTIN BIOSYNTHESIS
FILE REFERENCE: 3002-205

```

? CURRENT FILING DATE: 2005-08-09
? PRIOR APPLICATION NUMBER: US/06-07
? PRIOR FILING DATE: 2001-10-15
? PRIOR APPLICATION NUMBER: US 6
? PRIOR FILING DATE: 2000-10-13
? NUMBER OF SEQ ID NOS: 46
? SOFTWARE: PatentIn version 3.0.
? SEQ ID NO 15
? LENGTH: 8635

```



```
OY 310 SVG-----DIALPL-----HSVTPARPQEPVTLALERMALPG----- 341
;
;
Db 1745 ELGLSTLAVBLRNRLGARAETALPATLAFDHPTRPAIALDLLOAFSELAAAGTAAQA 1804
OY 342 PRSRALVTEGSAVV-----GIYTPSDVARLI-----PYRRAJOEP 377
;
;
Db 1805 PRARAHDEPRLIIVSMACRLPGGVDTPLALMQLLESGDAIGPFPEGMDVAGLYDPDP 1864
OY 378 TTTTSP-----ODADR 389
;
;
Db 1865 ---DAFGKSVTNLGGFLVDADR 1884

RESULT 7
US-11-087-099-10617
; Sequence 10617, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 10617
; LENGTH: 511
; TYPE: PRT
; ORGANISM: Torenia hybrida
US-11-087-099-10617

Query Match 5.6%; Score 112; DB 7; Length 511;
Best Local Similarity 24.9%; Pred. No. 0.32;
Matches 104; Conservative 48; Mismatches 146; Indels 120; Gaps 22;

OY 21 VLVTIMLFTMSLATV-----LPGTVGGYPAVYVWMLGAGAVMLLASLAHELAHAV 73
;
;
Db 19 VLVTIIVLSKSLSTIIVSRKRLPPGPGFPVY-----GALPILGSM--PHVALAKM 68
;
;
OY 74 ARRAG--VSBSVTMLFGVTLALGGA--KTPKAAF--RIAFAGPATSLASATFGALA 127
;
;
Db 69 AKKGVVWYLTGTLGTVVASTPSSARAFKLTLDNFSNRPNAG--ATHLA---YGAQD 123
;
;
OY 128 ITLA--GVTRPAIVSVAMWLTATVNLGLFVL--LPGAPLDGGLVYARLYMRKRGDSVR 183
;
;
Db 124 WVFAYGPR-----W---RLRLRLSNLHMLGKALDGMANVA--SELGHMLE 166
;
;
OY 184 AGIGARAGRYVALVIALGLAEFVAGLVGVMIAFGWIFFAAREETRISTQQLFA 243
;
;
Db 167 AMNRASRGEEV---VPEKLVYAMANNIGV---ILSRVFYTKGEE----- 208
;
;
OY 244 GVRVADAMTAQPHTAQWIVNEDFIQRYVLGERHSAYPV---ADRDGSIQGLVALROLRD 300
;
;
Db 209 -VNERKQVIVELMTSAGYFNIGDV-----PLAMMDLQGIERGMKALHKKFD 255
;
;
OY 301 -----VAPS-----RSTTSVGCIALPLHSVTPARQOE 328
;
;
Db 256 KLITRMLQHEASAKHPDLDAIIAASLDVSSSEERLSTSNIKALLNLPTAGDTSSS 315
;
;
OY 329 PLTLALERMALPGPRSRALVTEGSAVVG---LVTPSDVARLIDVYRLAOPEPTTSP 383
;
;
Db 316 TTEWALSEMLKKKGLKKAQOEEMDRVGRERLIVESDIEKL--GYLKAIKETEPRKH 371

RESULT 8
US-11-143-980-61
; Sequence 61, Application US/11143980
; Publication No. US20050272133A1
; GENERAL INFORMATION:
; APPLICANT: Hucul, John
; APPLICANT: He, Min
; APPLICANT: Halali, Bradley A.
; APPLICANT: Wagenaar, Melissa M.
; APPLICANT: Graziani, Edmund
```

```
; APPLICANT: Summers, Mia
; APPLICANT: Kuloweki, Kerry
; APPLICANT: Pong, Kevin
; TITLE OF INVENTION: Biosynthetic Gene Cluster for the Production of a Complex
; FILE REFERENCE: AM-101426US
; CURRENT APPLICATION NUMBER: US/11/143,980
; CURRENT FILING DATE: 2005-06-03
; PRIOR APPLICATION NUMBER: US 60/664,483
; PRIOR FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/576,895
; PRIOR FILING DATE: 2004-06-03
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 61
; LENGTH: 497
; TYPE: PRT
; ORGANISM: Streptomyces sp.
US-11-143-980-61

Query Match 5.6%; Score 110.5; DB 7; Length 497;
Best Local Similarity 24.2%; Pred. No. 0.41;
Matches 107; Conservative 40; Mismatches 144; Indels 151; Gaps 22;

OY 1 MRDAIPGRIAGFVNNHMSVTVLMLFTWSLATMLGTVGYPAV-----VY 48
;
;
Db 130 VRDHPSPRPG-----LIRMQTAAAMWILIGVGG--PLGGLLADHIGRWAF 177
;
;
OY 49 WL---LG-AGGAVMLLA-----SLAHELAHAV--VARRAGSVES 83
;
;
Db 178 WLNPLGLAGAVTVLPLPRRRPATPPSGRLDVAGIILLAGLALTLGSLKGNATRGH 237
;
;
OY 84 VTLW-----LFGVYAL-----GGAETPRAARIAPAGPATSLASATFGALAIT 129
;
;
Db 238 APSWTDPAVIGCLIGLALTLTLPVERRAAVPLPLFRHRTYTLTLAGFFQVAAA 297
;
;
OY 130 LAGRTPAIVSVAMWLTATVNLGLFNLPGAPLDGGLVYARLYMR--RQDSV----- 182
;
;
Db 298 PVGIFLPEYFGHIGHSATASGLL--LPLILGML--SNRLTAAVLSGHVKKPVLIGAG 355
;
;
OY 183 ---RAGIGARAGRYVALVIALGLAEFVAGLVGVMIAFGWIFFAAREETRIIS 237
;
;
Db 356 LITAGTAFAVLRITTPALTSVLLLVGLAGAGRMGGLTIA-----TQSA 401
;
;
OY 238 TQQLFAGRYVA--DAMTAQPHTAQWIVNEDFIQRYVLGERHSAYPAVDROGS-----I 289
;
;
Db 402 VPRADMGTATGSLTQOLGAGVLAGAQSIL-----GHSQAAPATAIGSTVSWSGA 456
;
;
OY 290 TGLVALQL---RVVARSRSSTTSVGDIALPLHSVPTA--RQEPPLALERMALGPRS 344
;
;
Db 457 AGLLALGALLMRDI-----SIATAGKRPAP----- 483
;
;
OY 345 RALVTEGSAVVGYTPSDVARL 366
;
;
Db 484 -----TSGTAV-----PAKADRL 496

RESULT 9
US-10-467-657-2362
; Sequence 2362, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASTIGNANI Vega
; APPLICANT: MONACT Elisabetha
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
```

NUMBER OF SEQ ID NOS: 9218
SOFTWARE: SeqWin99, version 1.04
SEQ ID NO: 2362
LENGTH: 488
TYPE: PRF
ORGANISM: Neisseria gonorrhoeae
US-10-467-657-2362

Query Match
Best Local Similarity 24.7%; Score 109; DB 6; Length 488;
Matches 78; Conservative 41; Mismatches 131; Indels 66; Gaps 16;

30 WSLATMLPCTVG-----YPAV---VYMLLGGAGVMT-----59
60 WA-ASILLPASAEPMLMMKPYSAVNSGDTAWMTAAVLMTLPGALFYGGMVRKKK 118
60 LASLAHBLAHAVVARRAGVSVEVTLMFGVTLALGSEAKT-----PKAARLAFAPR 113
119 LLSLTHHSGFSLTGLVILWAAV-GYSLATPGNAFISGLGRVPLSGMOTDAARMLTSP 177
114 ATSLASATFGALATTLACVTPAIVISVAMWL--ATVNLGLFNIL--PGA--PLDG 166
178 NAPTPEPEVFMFQMTFALITAILTGAFERMKYSAMMLFGIWFELVYVFGAHVWVG 237
167 GLVPAVYMRHGGDSVRAGIGARGRVVALVY--IALGLAEF---VAGLVGGWILA 219
238 GEMSKGGLVDYAGGV-VHINAGIAGLVAALVIGRIGYGRAMPHPHMMATLIGAMLM 296
220 FFGWTFFAAREERISTQQLFAGVRAVDAMTAQHTAPGHTINVEDFQIRVLEGRHA 279
297 F-GWGGFNG-----SALAADAAGAMAVTGVSAVGAAGWLACER-----IAGHKPSA 345
280 YPADRDSITGLVAL 295
346 LGLA--SGAVSGLVGI 359

RESULT 10
US-11-087-099-1493
Sequence 1493, Application US/11087099
Publication No. US20060041961A1
GENERAL INFORMATION:
APPLICANT: Adad, Mark S. et al.
TITLE OF INVENTION: Genes and Uses for Plant Improvement
FILE REFERENCE: 38-21(53450)B EP
CURRENT APPLICATION NUMBER: US/11/087,099
CURRENT FILING DATE: 2005-03-22
NUMBER OF SEQ ID NOS: 12464
SEQ ID NO 1493
LENGTH: 514
TYPE: PRF
ORGANISM: Thermobifida fusca
US-11-087-099-1493

Query Match
Best Local Similarity 21.2%; Score 109; DB 7; Length 514;
Matches 87; Conservative 46; Mismatches 126; Indels 152; Gaps 18;

44 PAVVYVWLGAGGAVNMLSLAHLAHAVVARRAGVSVEVTLMFG-----90
73 PAIVISFL--AAGAVCLAMCYAEFASIVPVASATVFGVATLGEFVAMWIGMDILBET 131
91 ---GVTALGGEAKTPKAPRIAPAGPATSLASATFGALATTLACVTPAIVISVAMWLA 147
132 LAASVVSVM-----EYAGDLALPTTVTLAGLPVILGAV-----167
148 TVNLLGLFNILFGAFLDGRVRAVLMRRHGDVRAAGIGARGRVVALVIALGLA--205
168 AIAVILGLVGM--GATLSG-----RVTAVVVALKVG--IVLFIIVGAAV 210
206 -----EFVAGLVGG---VMLATIGWFI 225
211 DEANWTPPIPAQPAADTTGVSDATLTHTVLFGLBETSPGWPVVAASVVEFSFI--FD 269

QY 226 FAAREESTRISTQQLFAGV-----RVADAMTA-QPH-----TAP-----259
DB 270 IVATTAESTRNPBRDIPVIGFISLLIYTLVMAVAAVTGMKPYELNVTAPLSDAFRSV 329
QY 260 -----GWINVEDFQIRVYTLGERHSAVPADDDGISTGLVALRQLRDVAPSRRS 307
DB 330 GADWAAATLISLGGVIGITTVILVLMGQARVAFMS--RDGLPLRALSRIPHFGTPVYTT 388
QY 308 TTSVGDIALPLHSVPTARPOEPLTALLERMAPLGRSPALVTEGSAVVGIV 358
DB 369 LITTAVALLGLVY-----ISTLEEMNIGTLFAFVVS-----VGV 427

RESULT 11
US-11-205-109-30
Sequence 30, Application US/11205109
Publication No. US20050287641A1
GENERAL INFORMATION:
APPLICANT: Farnet, Chris
APPLICANT: Zazopoulos, Emmanuel
APPLICANT: Staffa, Alfredo
TITLE OF INVENTION: GENE CLUSTER FOR RAMOPLANIN BIOSYNTHESIS
FILE REFERENCE: 3002-2US
CURRENT APPLICATION NUMBER: US/11/205,109
CURRENT FILING DATE: 2005-08-17
PRIOR APPLICATION NUMBER: US/09/976,059
PRIOR FILING DATE: 2001-10-15
PRIOR APPLICATION NUMBER: US 60/239,924
PRIOR FILING DATE: 2000-10-13
NUMBER OF SEQ ID NOS: 46
SOFTWARE: PatentIn version 3.0
SEQ ID NO 30
LENGTH: 619
TYPE: PRF
ORGANISM: Actinoplanes sp.
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)..(1)
OTHER INFORMATION: V represents a non-standard initiator codon. It is expected that
OTHER INFORMATION: at this position
US-11-205-109-30

Query Match
Best Local Similarity 31.4%; Score 109; DB 7; Length 619;
Matches 82; Conservative 22; Mismatches 93; Indels 64; Gaps 19;

QY 4 AIPGRIAGFVNNVMSVYLIMLFTWSLATMLPCTVG-----YPAVYVWLGAGGA 56
DB 85 SITDKLAGAFVPOALSARV-FGFHQMSLA--LPQAVEGVIALVLYRAVRRW--HGRGA 139
QY 57 VMLASLAHLAHAVVARRAGVSVE--SVTLMFGVYVTLALGGEAKTPKAPRIAPAP 114
DB 140 GLAAGLFA---TPPIVSMGSHMEDGALTCLVLAADAFGA-AVTGSPARLALAGW 195
QY 115 TSLASLA-----TFGLAIT-LAG--VTPAIVISVAMWLA--TVNLLGL-FNLLP 160
DB 196 IGLGQAMQAMWLPLVTVTLAGAFVRAARVAVAAVAATLAVSLMLVLTLLP 255
QY 161 GAP--LDGRLVRAV--LMRRHGDVRAAGI--GAARAGVVALVIALGLAEFVAG--211
DB 266 GSHRPWADGTTSGNAFAVWFGVNGFD--RAGIHVFGALTTG-----FTDGA 300
QY 212 LVGVWLA-----IGWF 224
DB 301 AAGGSWTALAADRLATQIGW 321

RESULT 12
US-10-858-730-121
Sequence 121, Application US/10858730
Publication No. US20050255568A1

GENERAL INFORMATION:
; APPLICANT: Bailey, Richard B.
; APPLICANT: Blomquist, Paul
; APPLICANT: Doenen, Reed
; APPLICANT: Driggers, Edward M.
; APPLICANT: Madden, Kevin T.
; APPLICANT: O'Leary, Jessica
; APPLICANT: O'Toole, George
; APPLICANT: Trueheart, Joshua
; APPLICANT: Walbridge, Michael J.
; APPLICANT: Yorgey, Peter S.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR AMINO ACID
; FILE REFERENCE: 14184-030001
; CURRENT APPLICATION NUMBER: US/10/858,730
; CURRENT FILING DATE: 2004-06-01
; PRIOR APPLICATION NUMBER: US 60/475,000
; PRIOR FILING DATE: 2003-05-30
; PRIOR APPLICATION NUMBER: US 60/551,860
; PRIOR FILING DATE: 2004-03-10
; NUMBER OF SEQ ID NOS: 364
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 121
; LENGTH: 334
; TYPE: PRT
; ORGANISM: Streptomyces coelicolor
US-10-858-730-121

Query Match 5.4%; Score 106.5; DB 6; Length 334;
Best Local Similarity 24.3%; Pred. No. 0.51;
Matches 81; Conservative 32; Mismatches 111; Indels 109; Gaps 16;

QY 2 RDA--IPGRI--AGFVNVMVSVLVILMLFTWSLATMLPGTVGVPAVVYMLLAGGA 56
DB 61 RDAGQIAGLSGSGYLSVYATL-----GVSSGTALTI-----DGV 100
QY 57 VMLASLAHSLAHAAVAR-----RAGVSVSVTLMLF----- 89
DB 101 QPLVAGALAGPLIRGYVRGQWGLMLGSGVATVTVDAGAAAGAVMMATLVPLGML 160
QY 90 --GGTALGGEAKTKAFRIAF--GPAISLSTFGALATTAGVTPAIVISVAMWL 146
DB 161 SLVAATFLEGRTVRVAP--RVALLTHCATSAVL--FSGLAGLGAAPPA--GSSFWL 213
QY 147 ATVNLLGLFNLPGAPLDGRLVRAVLMRRHGS-----VRAGIGAPARARHV 195
DB 214 ATAMLVV-----LPTFGYGLVWLIIRSGITREVTMLMAPVAVMGALMFGSPF 265
QY 196 ALVLIALGLAEFVAGLV--GGVWLAFIGWFIFAAREBEETRISTQULPAGVRVADAMT 252
DB 266 G-VQPALGLAVGLAAVVVVRRGG-----GARRRPVRSAGDRPAAGGPTADQPT 313
QY 253 AQPHTAPGMINVEDFIQRYVLGERHSAYPADR 285
DB 314 NRPTDRPTAAGSTD-----RPTADR 333

RESULT 13
US-11-087-099-6812
; Sequence 6812, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 6812
; LENGTH: 434
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis H37Rv
US-11-087-099-6812

Query Match 5.3%; Score 105.5; DB 7; Length 434;
Best Local Similarity 22.8%; Pred. No. 0.83;
Matches 89; Conservative 39; Mismatches 137; Indels 125; Gaps 19;
QY 7 LGRAGFVNVMHMSVLVILMLFTWSLATMLPGTVGVPAVVYMLLAGGAVMLLASLAAH 66
DB 66 LGGWAGF--SVGM--LYWTF--WVIYVGEBAVAG--KVLTYWYI----- 101
QY 67 ELAAVVARRAGVSVESVTLMLFGGYTALGGEAKTPAAAFRIAPAGPATSLASATGAL 126
DB 102 -----DAPLML-----ASJCLMMMTATNLVSVSSGFEF 130
QY 127 AITLAGVTPAIVISVAMMLATVNLIG--LENLPGAPLDGRLVRAVLMRRHGSVR 183
DB 131 EFWFAGVYVATIV-----GELVIGTAFAFGLLGHGHDFFSNL-----SAHGCFEP 175
QY 184 AGTGAARAGRVVALVLIALGLAEFVAGLVGVGWLAIQWFIFAAREBEETRISTQULFA 243
DB 176 DGVGAVPAAIYVA--IFSMGTGEVVT-----IAAARAPDQRAVQRAWS 217
QY 244 GV--RVADAMTAQHHTAP--GWINVEDFIQRYVLGERHSAYPADRDGSIYGLVALROL 298
DB 218 TVVARIYIFFVGSVFLTLVILPMNSLEIGASPYVAALRHMGIGAGDQ---IMNAVULTAV 274
QY 299 RDVAPSRRTSVGDIALPLHSVTPARPOEBTLALERMALPGRSALVTESAVVGI- 357
DB 275 LSCINSGLYTAS-----RMLFVLAARQEAR--AQLVYVNRGVPTPAIM--GSSVVGFL 324
QY 358 -----VTPSDV-----ARLIDYRL 372
DB 325 CVIMAWVSPATVFEVFLNLSGAVILFVYL 354

RESULT 14
US-11-087-099-3033
; Sequence 3033, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 3033
; LENGTH: 488
; TYPE: PRT
; ORGANISM: Rhodococcus erythropolis
US-11-087-099-3033

Query Match 5.3%; Score 104.5; DB 7; Length 488;
Best Local Similarity 24.3%; Pred. No. 1.1;
Matches 56; Conservative 25; Mismatches 62; Indels 87; Gaps 12;

QY 32 LATMLPGTVGYP---AVVYMLLAGGAVMLLASLAHSLAHAAVARRAVSVSVYTL- 86
DB 96 MSSALPAAAGGYTPARABMPW---GSPATGTAILIEYSIAPAAIATFGAYVESINLF 151
QY 87 -----WLFQGYTALGGEAKTKAFRIAPAG--PATSLASTFGALATTAGVTPA 137
DB 152 GITDGMWVYLAIVAI-----FTGHLTGAGELKMKMFIITTAIVGL----- 193
QY 138 IIVISVAMMLATVNLIGFNLPGAPLDGRLVRAVLMRRHGSVRAIGIAGARGVVAL 197
DB 194 VWFVVS-----AVGLF-----DSSULTD-----IAV 214
QY 198 VLIALGLAEFVAGLVGVGWLDA--FIGWFIFAA-----AREEETRI 236
DB 215 DTSVAGSSSFLPFGLL-GVMAAVPFAIMFILAEGVPLAAEAREPEKGV 263

RESULT 15

```

US-11-087099-9168
; Sequence 9168. Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450) B EP
; CURRENT APPLICATION NUMBER: US/11/087, 099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 9168
; LENGTH: 557
; TYPE: PRN
; ORGANISM: Thermoplasma acidophilum
US-11-087-099-9168

```

Query Match	5.3%;	Score 104.5;	DB 7;	length 557;
Best Local Similarity	21.0%;	Pred. No. 1.3;		
Matches	61;	Mismatches	86;	Indels 103; Gaps 13,

Qy 26 MLEFMSLTMLTGGTGGYPAVVYWMLLGAGAVMLSLAHLEIHLNV-----VARRAGS 80

Dh 39 MLFAAGASMLAGPA-----AYLSMTII--GGILYLVLTALVYAEIAGCMIPRGALVRRGQYS 92

Qy 81 VESVTLTMLFG-----GVTAIAGEAKTPPKAFAPIAAGPATSLALSATGALAITLAG 132

Dh 93 HGGLGLPFLFGMAVYFLSNAVSVPALIEAGV-----ITYAGTYVKGAVTPR--GVL--TAVG 142

Qy 133 VRTPAIVL-----SVAMW-----LATVNNLLGVPNLLPGA 162

Dh 143 ILLAAILMLGPPFLAYFGIKVMGKNTGTMIMWLLIIPSATVALLLVHFNIGNPFELKTGF 202

Qy 163 PLDGG-----GRLYRAYILMRR-----HGDSVRAGIGAAAGRVAVLV----- 199

Dh 203 VTYGMSVFEAISGIVESYLGFRQALDYGGAKAPQRSVPATATLSVIGIALVALTQ 262

Qy 200 -IAGLAERVAGLYGVW-----LAIGWPIFAA 229

Dh 263 VVFIQVNMVAAGVAPGQMSALSGGVYTAAPATLASSAGVLFVSTYSLFPA 313

Search completed: March 23, 2006, 06:29:39
Job time : 11.431 secs

This Page Blank (usp.)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using SW model

Run on: March 23, 2006, 04:56:29 ; Search time 109.803 Seconds
(without alignments)
1652.628 Million cell updates/sec

Title: US-10-617-038-29

Perfect score: 2195
Sequence: 1 MASSASDCTHERSAFRLSP.....AALEQSGILDAPRTGRDRA 413

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_21:*

- 1: geneeqp1980s:*
- 2: geneeqp1990s:*
- 3: geneeqp2000s:*
- 4: geneeqp2001s:*
- 5: geneeqp2002s:*
- 6: geneeqp2003s:*
- 7: geneeqp2003bs:*
- 8: geneeqp2004s:*
- 9: geneeqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2195	100.0	413	AD137308	Ad137308 M. tuberc
2	149	6.7	346	AAg20855	AAg20855 Arabidops
3	148	6.7	353	AAg20854	AAg20854 Arabidops
4	148	6.7	356	AAg20853	AAg20853 Arabidops
5	146.5	6.7	369	ADx77895	ADx77895 Plant ful
6	146.5	6.7	369	ADx68280	ADx68280 Plant ful
7	130	5.9	172	ABM92392	ABM92392 M. xanthu
8	126	5.7	582	ADCS2136	ADCS2136 M. xanthu
9	121	5.5	338	ABM67066	ABM67066 Drosophi
10	118.5	5.4	322	ADT59525	ADT59525 Plant pol
11	117.5	5.4	889	ABO60724	ABO60724 Klebsiell
12	112.5	5.1	409	ABO71490	ABO71490 Pseudomon
13	112	5.1	574	ABU34102	ABU34102 Protein e
14	109.5	5.0	1127	ABO78103	ABO78103 Pseudomon
15	106.5	4.9	691	ABG61612	ABG61612 Human DPR
16	106	4.8	452	ADY04294	ADY04294 Plant ful
17	105	4.8	343	AAU49953	AAU49953 Propionib
18	105	4.8	343	ABM46472	ABM46472 Propionib
19	104.5	4.8	742	AD527327	AD527327 Bacterial
20	104.5	4.8	742	AD526598	AD526598 Bacterial
21	104.5	4.8	742	AD526973	AD526973 Bacterial
22	102.5	4.7	4635	ADK56098	ADK56098 Streptomy
23	102	4.6	563	ABO68803	ABO68803 Pseudomon
24	101	4.6	518	ADT50839	ADT50839 Cancer re

25	100.5	4.6	586	9	AEb03469	AEb03469 Mycobacte
26	100.5	4.6	586	9	AEa79372	AEa79372 Novel M.
27	100	4.6	421	8	AD595064	AD595064 Human the
28	100	4.6	421	8	ADY67921	ADY67921 Biologica
29	99.5	4.5	424	8	ADY08484	ADY08484 Plant ful
30	99.5	4.5	981	9	ABM97613	ABM97613 M. xanthu
31	99	4.5	421	8	ADN61819	ADN61819 Human nov
32	98.5	4.5	748	9	ABM94678	ABM94678 M. xanthu
33	98.5	4.5	1178	9	ABM96130	ABM96130 M. xanthu
34	98	4.5	668	4	AAb79081	AAb79081 Corynebac
35	98	4.5	706	4	AAAG90125	AAAG90125 C. glutam
36	98	4.5	706	7	ADP13214	ADP13214 C. glutam
37	98	4.5	706	7	ADP13657	ADP13657 C. glutam
38	97	4.4	401	3	AAV77188	AAV77188 S. venez
39	97	4.4	401	3	AAV80998	AAV80998 S. venez
40	97	4.4	402	5	AAE24236	AAE24236 Streptomy
41	97	4.4	402	8	ADL91909	ADL91909 Streptomy
42	97	4.4	483	4	AAAB93479	AAAB93479 Human pro
43	97	4.4	3782	3	AAV77179	AAV77179 S. venez
44	97	4.4	3782	5	AAE24228	AAE24228 Streptomy
45	97	4.4	3782	8	ADL91932	ADL91932 Streptomy

ALIGNMENTS

RESULT 1	AD137308	standard: protein; 413 AA.
ID	AD137308	
AC	AD137308;	
XX		
DT	22-APR-2004 (first entry)	
XX		
DE	M. tuberculosis low oxygen induced antigen Rv2727c SEQ ID NO:29.	
XX		
KW	mycobacterial infection; vaccine; tuberculosis;	
KW	Mycobacterium tuberculosis; immunisation; antibacterial; gene therapy;	
KW	low oxygen induced antigen.	
XX		
OS	Mycobacterium tuberculosis.	
XX		
PN	WO2004006952-A2.	
XX		
PD	22-JAN-2004.	
XX		
PF	08-JUL-2003; 2003WO-DK000477.	
XX		
PR	13-JUL-2002; 2002DK-00001098.	
XX		
PA	(STAT-) STATENS SERUM INST.	
XX		
PI	Andersen P, Rosenkrands I, Stryhn A;	
XX		
DR	WPI; 2004-122778/12.	
XX		
DR	N-PSDB; AD137353.	
XX		
PT	Use of one or more polypeptides or their fragments, which are expressed	
PT	during the latent stage of the mycobacterial infection, and/or nucleic	
PT	acids encoding the polypeptides, for a therapeutic vaccine against	
PT	tuberculosis.	
XX		
PS	Claim 3; SEQ ID NO 29; 76pp; English.	
XX		
CC	The present invention describes polypeptides or their fragments, which	
CC	are expressed during the latent stage of a mycobacterial infection,	
CC	and/or nucleic acids encoding the polypeptides, which are useful for	
CC	creating a therapeutic vaccine against tuberculosis. Also described: (1)	
CC	a therapeutic vaccine against tuberculosis comprising one or more	
CC	polypeptides; (2) a method for treating an animal, including a human	
CC	being, with tuberculosis caused by virulent mycobacteria, e.g. by	
CC	Mycobacterium tuberculosis, M. africanum or M. bovis; (3) a method for	
CC	immunising an animal, including a human being, against tuberculosis	

CC caused by virulent mycobacteria; (4) a method of diagnosing tuberculosis
CC caused by virulent mycobacteria in an animal, including a human being;
CC (5) a method for diagnosing previous or ongoing infection with a virulent
CC mycobacterium; and (6) a method of diagnosing Mycobacterium tuberculosis
CC infection in a subject. The polypeptides have antibacterial activities,
CC and can be used in vaccines and in gene therapy. The polypeptides are
CC useful for the manufacture of a therapeutic vaccine for treating an
CC individual who is infected by a virulent mycobacterium, e.g. M.
CC tuberculosis, and who is not vaccinated with BCG against tuberculosis.
CC The present sequence represents a low oxygen induced antigen, which is
CC used in the exemplification of the present invention.

XX Sequence 413 AA;

Query Match 100.0%; Score 2195; DB 8; Length 413;

Best Local Similarity 100.0%; Pred. No. 1.8e-220; Mismatches 0; Indels 0; Gaps 0;

Matches 413; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASSASDGTHERSAFRLSPVLGSGMGPMTGLVVAQSWRDYLGQPDKLPAPRTIAL 60

Db 1 MASSASDGTHERSAFRLSPVLGSGMGPMTGLVVAQSWRDYLGQPDKLPAPRTIAL 60

QY 61 AAQAFRDIVILGKARRPVSNHRYFERISQEVAAGLFYGRRWLEKPSGFAQPPPLT 120

Db 61 AAQAFRDIVILGKARRPVSNHRYFERISQEVAAGLFYGRRWLEKPSGFAQPPPLT 120

QY 121 EVAARKVDRRRSFYRIFFDSGFTPHPEPGSQRMISTYANNREYALLRRHPEPPWLYC 180

Db 121 EVAARKVDRRRSFYRIFFDSGFTPHPEPGSQRMISTYANNREYALLRRHPEPPWLYC 180

QY 121 EVAARKVDRRRSFYRIFFDSGFTPHPEPGSQRMISTYANNREYALLRRHPEPPWLYC 180

Db 121 EVAARKVDRRRSFYRIFFDSGFTPHPEPGSQRMISTYANNREYALLRRHPEPPWLYC 180

QY 181 VHGTMGRAPIDLAVFRAMKLDDELGLNIWMPVLPMDGPRGQGLPKGAVFGEEDYLDVH 240

Db 181 VHGTMGRAPIDLAVFRAMKLDDELGLNIWMPVLPMDGPRGQGLPKGAVFGEEDYLDVH 240

QY 241 GTAQAVMIRRLLSMIRSOEESLIGLNGLSIGYITASLVASLEKGLACALLGVVADLI 300

Db 241 GTAQAVMIRRLLSMIRSOEESLIGLNGLSIGYITASLVASLEKGLACALLGVVADLI 300

QY 301 ELLGRHCGLRHDPFRHTVKNAPIGRMISPLSTPLVPMGPFYAGIADRLVHPRROV 360

Db 301 ELLGRHCGLRHDPFRHTVKNAPIGRMISPLSTPLVPMGPFYAGIADRLVHPRROV 360

QY 361 TRLMEHWCKPEIWMYPGHGTGFGQSRPVRRFYQALLDSGLIDAPRTORDSA 413

Db 361 TRLMEHWCKPEIWMYPGHGTGFGQSRPVRRFYQALLDSGLIDAPRTORDSA 413

QY 361 TRLMEHWCKPEIWMYPGHGTGFGQSRPVRRFYQALLDSGLIDAPRTORDSA 413

Db 361 TRLMEHWCKPEIWMYPGHGTGFGQSRPVRRFYQALLDSGLIDAPRTORDSA 413

QY 361 TRLMEHWCKPEIWMYPGHGTGFGQSRPVRRFYQALLDSGLIDAPRTORDSA 413

Db 361 TRLMEHWCKPEIWMYPGHGTGFGQSRPVRRFYQALLDSGLIDAPRTORDSA 413

QY 361 TRLMEHWCKPEIWMYPGHGTGFGQSRPVRRFYQALLDSGLIDAPRTORDSA 413

Db 361 TRLMEHWCKPEIWMYPGHGTGFGQSRPVRRFYQALLDSGLIDAPRTORDSA 413

QY 361 TRLMEHWCKPEIWMYPGHGTGFGQSRPVRRFYQALLDSGLIDAPRTORDSA 413

Db 361 TRLMEHWCKPEIWMYPGHGTGFGQSRPVRRFYQALLDSGLIDAPRTORDSA 413

QY 361 TRLMEHWCKPEIWMYPGHGTGFGQSRPVRRFYQALLDSGLIDAPRTORDSA 413

Db 361 TRLMEHWCKPEIWMYPGHGTGFGQSRPVRRFYQALLDSGLIDAPRTORDSA 413

QY 361 TRLMEHWCKPEIWMYPGHGTGFGQSRPVRRFYQALLDSGLIDAPRTORDSA 413

Db 361 TRLMEHWCKPEIWMYPGHGTGFGQSRPVRRFYQALLDSGLIDAPRTORDSA 413

QY 361 TRLMEHWCKPEIWMYPGHGTGFGQSRPVRRFYQALLDSGLIDAPRTORDSA 413

Db 361 TRLMEHWCKPEIWMYPGHGTGFGQSRPVRRFYQALLDSGLIDAPRTORDSA 413

PR 25-MAR-1999; 99US-0126264P.

PR 29-MAR-1999; 99US-0126785P.

PR 01-APR-1999; 99US-0127462P.

PR 06-APR-1999; 99US-0128234P.

PR 08-APR-1999; 99US-0128714P.

PR 16-APR-1999; 99US-0129845P.

PR 19-APR-1999; 99US-0130077P.

PR 21-APR-1999; 99US-0130449P.

PR 23-APR-1999; 99US-0130510P.

PR 28-APR-1999; 99US-0130891P.

PR 30-APR-1999; 99US-0132048P.

PR 30-APR-1999; 99US-0132407P.

PR 04-MAY-1999; 99US-0132484P.

PR 05-MAY-1999; 99US-0132485P.

PR 06-MAY-1999; 99US-0132486P.

PR 06-MAY-1999; 99US-0132487P.

PR 07-MAY-1999; 99US-0132863P.

PR 11-MAY-1999; 99US-0134256P.

PR 14-MAY-1999; 99US-0134219P.

PR 14-MAY-1999; 99US-0134221P.

PR 14-MAY-1999; 99US-0134370P.

PR 18-MAY-1999; 99US-0134768P.

PR 19-MAY-1999; 99US-0134941P.

PR 20-MAY-1999; 99US-0135124P.

PR 21-MAY-1999; 99US-0135353P.

PR 24-MAY-1999; 99US-0135629P.

PR 25-MAY-1999; 99US-0136021P.

PR 27-MAY-1999; 99US-0136392P.

PR 28-MAY-1999; 99US-0136782P.

PR 01-JUN-1999; 99US-0137222P.

PR 03-JUN-1999; 99US-0137528P.

PR 04-JUN-1999; 99US-0137502P.

PR 07-JUN-1999; 99US-0137724P.

PR 08-JUN-1999; 99US-0138094P.

PR 10-JUN-1999; 99US-0138540P.

PR 10-JUN-1999; 99US-0138847P.

PR 14-JUN-1999; 99US-0139119P.

PR 16-JUN-1999; 99US-0139452P.

PR 17-JUN-1999; 99US-0139492P.

PR 18-JUN-1999; 99US-0139454P.

PR 18-JUN-1999; 99US-0139455P.

PR 18-JUN-1999; 99US-0139456P.

PR 18-JUN-1999; 99US-0139457P.

PR 18-JUN-1999; 99US-0139458P.

PR 18-JUN-1999; 99US-0139459P.

PR 18-JUN-1999; 99US-0139460P.

PR 18-JUN-1999; 99US-0139461P.

PR 18-JUN-1999; 99US-0139462P.

PR 18-JUN-1999; 99US-0139463P.

PR 18-JUN-1999; 99US-0139750P.

PR 18-JUN-1999; 99US-0139763P.

PR 21-JUN-1999; 99US-0139817P.

PR 22-JUN-1999; 99US-0139899P.

PR 23-JUN-1999; 99US-0140353P.

PR 23-JUN-1999; 99US-0140354P.

PR 24-JUN-1999; 99US-0140695P.

PR 28-JUN-1999; 99US-0140823P.

PR 29-JUN-1999; 99US-0140951P.

PR 30-JUN-1999; 99US-0141267P.

PR 01-JUL-1999; 99US-0141842P.

PR 01-JUL-1999; 99US-0142154P.

PR 02-JUL-1999; 99US-0142055P.

PR 06-JUL-1999; 99US-0142390P.

PR 08-JUL-1999; 99US-0142803P.

PR 09-JUL-1999; 99US-0142920P.

PR 12-JUL-1999; 99US-0142977P.

PR 13-JUL-1999; 99US-0143542P.

PR 14-JUL-1999; 99US-0143624P.

PR 15-JUL-1999; 99US-0144005P.

PR 16-JUL-1999; 99US-0144085P.

PR	08-OCT-1999;	99US-0158232P.
PR	12-OCT-1999;	99US-0158366P.
PR	13-OCT-1999;	99US-0159293P.
PR	13-OCT-1999;	99US-0159294P.
PR	13-OCT-1999;	99US-0159295P.
PR	14-OCT-1999;	99US-0159329P.
PR	14-OCT-1999;	99US-0159330P.
PR	14-OCT-1999;	99US-0159331P.
PR	14-OCT-1999;	99US-0159637P.
PR	14-OCT-1999;	99US-0159638P.
PR	18-OCT-1999;	99US-0159584P.
PR	21-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.
PR	21-OCT-1999;	99US-0160768P.
PR	21-OCT-1999;	99US-0160770P.
PR	21-OCT-1999;	99US-0160814P.
PR	21-OCT-1999;	99US-0160815P.
PR	22-OCT-1999;	99US-0160980P.
PR	22-OCT-1999;	99US-0160981P.
PR	22-OCT-1999;	99US-0160989P.
PR	25-OCT-1999;	99US-0161404P.
PR	25-OCT-1999;	99US-0161405P.
PR	25-OCT-1999;	99US-0161406P.
PR	26-OCT-1999;	99US-0161359P.
PR	26-OCT-1999;	99US-0161360P.
PR	26-OCT-1999;	99US-0161361P.
PR	28-OCT-1999;	99US-0161920P.
PR	28-OCT-1999;	99US-0161992P.
PR	28-OCT-1999;	99US-0161993P.
PR	29-OCT-1999;	99US-0162142P.

Query Match	6.7%;	Score 148;	DB 3;	Length 346;
Best Local Similarity	27.7%;	Pred. No. 3e-06;		
Matches	93;	Conservative	38;	Mismatches 133;
				Indels 72;
				Gaps 20;

```

QY      82 NHRVFERISQVAAGLEFYGNRRMLERKSGFFAAPPRLTIVAKKIDRRSFFR--1FF 133
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      32 NLELLERVQGLFP-LEVOG-QNN-----PPPLVRVMTWETKATLREGVF- 78
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      140 DSGFLPHGE-----PSQRWLSYTANNREYALLRHEPRPWLVCYHGTGEMGRAPLDLA 194
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      79 ---QTPCADELTAALPPESR---TA-RVAMLPVKVPPQKMACVNHLAGTGHTYD- 127
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      195 VFRAWKLDHEL-GNIVMPVL-PMHGPRGQGLRKA-VFGEEDVLDDVHSTAQVMDIR 250
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      128 --RRLRGGPLVKONITATWLESFFYGQRRPFLQGRALICVSDLL-LLGRA-TIESR 182
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      251 RLTSWIRSQEESLIGNGSLGCIYASLVALEBGLACALLGVVDLLELLGRH---- 306
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      183 SLIIMLDTEEBFGKMGVCGLSMGVSHMSQSLRP-----TPVATL-PFLSPHSAVV 233
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      307 --C-----GLRHK-DPRHIVKMAEPICRMISPLSLT-----PLVPMGRFLYA 347
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      234 AFGGILKYGTAMEALREBELLAQKITMTLDEVERMERENVISLTDTVTRFPIKNDVAVIFV 293
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      348 GIADRLVHPREQVTRLMEHMKPEITVYPGHTGFF 383
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      294 AATDGVLPKHSVLELQKAPGSEVRAVWTGGHVSF 329
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 3
AAG20854
ID AAG20854 standard; protein; 353 AA.
XX
XX
AC AAG20854;
XX
XX
DT 17-OCT-2000 (first entry)
XX
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 23198.
XX
KW protein identification; signal transduction pathway; metabolic pathway;
KW hydridsation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

```

XX	Arabidopsis thaliana.	PR	29-JUN-1999;	99US-0140991P.
OS		PR	30-JUN-1999;	99US-0141287P.
XX	EPI033405-A2.	PR	01-JUL-1999;	99US-0141842P.
PN		PR	01-JUL-1999;	99US-0142154P.
XX		PR	02-JUL-1999;	99US-0142055P.
PD	06-SEP-2000.	PR	06-JUL-1999;	99US-0142390P.
XX		PR	08-JUL-1999;	99US-0142803P.
XX		PR	09-JUL-1999;	99US-0142920P.
PF	25-FEB-2000; 2000EP-00301439.	PR	12-JUL-1999;	99US-0142977P.
XX		PR	13-JUL-1999;	99US-0143542P.
XX		PR	14-JUL-1999;	99US-0143624P.
PR	25-FEB-1999;	PR	15-JUL-1999;	99US-0144005P.
PR	05-MAR-1999;	PR	16-JUL-1999;	99US-0144005P.
PR	09-MAR-1999;	PR	16-JUL-1999;	99US-0144086P.
PR	23-MAR-1999;	PR	16-JUL-1999;	99US-0144325P.
PR	25-MAR-1999;	PR	19-JUL-1999;	99US-0144331P.
PR	29-MAR-1999;	PR	19-JUL-1999;	99US-0144332P.
PR	01-APR-1999;	PR	19-JUL-1999;	99US-0144333P.
PR	06-APR-1999;	PR	19-JUL-1999;	99US-0144333P.
PR	08-APR-1999;	PR	19-JUL-1999;	99US-0144334P.
PR	16-APR-1999;	PR	19-JUL-1999;	99US-0144335P.
PR	19-APR-1999;	PR	20-JUL-1999;	99US-0144352P.
PR	21-APR-1999;	PR	20-JUL-1999;	99US-0144632P.
PR	23-APR-1999;	PR	20-JUL-1999;	99US-0144684P.
PR	23-APR-1999;	PR	21-JUL-1999;	99US-0144814P.
PR	28-APR-1999;	PR	21-JUL-1999;	99US-0145086P.
PR	30-APR-1999;	PR	21-JUL-1999;	99US-0145086P.
PR	04-MAY-1999;	PR	22-JUL-1999;	99US-0145087P.
PR	05-MAY-1999;	PR	22-JUL-1999;	99US-0145089P.
PR	06-MAY-1999;	PR	22-JUL-1999;	99US-0145182P.
PR	06-MAY-1999;	PR	22-JUL-1999;	99US-0145182P.
PR	07-MAY-1999;	PR	23-JUL-1999;	99US-0145145P.
PR	11-MAY-1999;	PR	23-JUL-1999;	99US-0145218P.
PR	14-MAY-1999;	PR	26-JUL-1999;	99US-0145224P.
PR	14-MAY-1999;	PR	26-JUL-1999;	99US-0145276P.
PR	14-MAY-1999;	PR	27-JUL-1999;	99US-0145913P.
PR	14-MAY-1999;	PR	27-JUL-1999;	99US-0145918P.
PR	18-MAY-1999;	PR	27-JUL-1999;	99US-0145919P.
PR	19-MAY-1999;	PR	28-JUL-1999;	99US-0145951P.
PR	20-MAY-1999;	PR	02-AUG-1999;	99US-0146386P.
PR	21-MAY-1999;	PR	02-AUG-1999;	99US-0146388P.
PR	24-MAY-1999;	PR	02-AUG-1999;	99US-0146389P.
PR	25-MAY-1999;	PR	03-AUG-1999;	99US-0147038P.
PR	27-MAY-1999;	PR	04-AUG-1999;	99US-0147204P.
PR	28-MAY-1999;	PR	04-AUG-1999;	99US-0147302P.
PR	01-JUN-1999;	PR	05-AUG-1999;	99US-0147192P.
PR	03-JUN-1999;	PR	05-AUG-1999;	99US-0147260P.
PR	04-JUN-1999;	PR	06-AUG-1999;	99US-0147303P.
PR	07-JUN-1999;	PR	06-AUG-1999;	99US-0147416P.
PR	08-JUN-1999;	PR	09-AUG-1999;	99US-0147453P.
PR	10-JUN-1999;	PR	09-AUG-1999;	99US-0147935P.
PR	10-JUN-1999;	PR	10-AUG-1999;	99US-0148171P.
PR	14-JUN-1999;	PR	11-AUG-1999;	99US-0148319P.
PR	16-JUN-1999;	PR	12-AUG-1999;	99US-0148341P.
PR	16-JUN-1999;	PR	13-AUG-1999;	99US-0148555P.
PR	17-JUN-1999;	PR	13-AUG-1999;	99US-0148684P.
PR	18-JUN-1999;	PR	16-AUG-1999;	99US-0149368P.
PR	18-JUN-1999;	PR	17-AUG-1999;	99US-0149175P.
PR	18-JUN-1999;	PR	18-AUG-1999;	99US-0149426P.
PR	18-JUN-1999;	PR	20-AUG-1999;	99US-0149722P.
PR	18-JUN-1999;	PR	20-AUG-1999;	99US-0149723P.
PR	18-JUN-1999;	PR	20-AUG-1999;	99US-0149929P.
PR	18-JUN-1999;	PR	23-AUG-1999;	99US-0149902P.
PR	18-JUN-1999;	PR	23-AUG-1999;	99US-0149930P.
PR	18-JUN-1999;	PR	25-AUG-1999;	99US-0150566P.
PR	18-JUN-1999;	PR	26-AUG-1999;	99US-0150884P.
PR	18-JUN-1999;	PR	27-AUG-1999;	99US-0151065P.
PR	18-JUN-1999;	PR	27-AUG-1999;	99US-0151066P.
PR	21-JUN-1999;	PR	27-AUG-1999;	99US-0151080P.
PR	22-JUN-1999;	PR	30-AUG-1999;	99US-0151303P.
PR	23-JUN-1999;	PR	31-AUG-1999;	99US-0151438P.
PR	23-JUN-1999;	PR	01-SEP-1999;	99US-0151930P.
PR	24-JUN-1999;	PR	07-SEP-1999;	99US-0152363P.
PR	28-JUN-1999;	PR	10-SEP-1999;	99US-0153070P.

PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156596P.
PR 29-SEP-1999; 99US-0157117P.
PR 04-OCT-1999; 99US-0157753P.
PR 05-OCT-1999; 99US-0157865P.
PR 06-OCT-1999; 99US-0158029P.
PR 07-OCT-1999; 99US-0158332P.
PR 08-OCT-1999; 99US-0158363P.
PR 12-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 6.7%; Score 148; DB 3; Length 353;
Beet Local Similarity 27.7%; Pred. No. 3,1e-06;

Matches 93; Conservative 38; Mismatches 133; Indels 72; Gaps 20;

QY 82 NHRVERISQEVAGLEFYGNRRMLEKPSGFPAOPPLEVAVRKVDRRSPYR--TF 139
DB 39 NLELLERAVQRLFP-LEVOG-QNW-----PPLVPRPVKRTVMETKTATLREGVF- 85
QY 140 DSGFTPHDE-----PGSQRWLSTYANNREYALLLRHDEPRMVLVCVHTEMGRAPLDLA 194
DB 86 ---GTPCADLITLALPPESR-----TA--RVAMLVKQVPPQMACVHLAAGDHTYD-- 134
QY 195 VFRAKMLDEL-GLNTYMPVL--PMAGPRGGLPFGA-VFPGEDVLDDVHAGTAQAVMIR 250
DB 135 --RRRLRGGLPVKQNIATMVLSPFYGQRPRPLQCGARLLCVSDLL--LLGGA-TIESR 189
QY 251 RLWSIRQOEBSLTLGLNGLSIGYIASIVASLEBGLACALIGVPADLIELGRH---- 306
DB 190 SLIHMLDTBEGKMGVCGLSGVHASMVCSLHP-----TVATL-PLLSHSAV 240
QY 307 --C-----GLRHK-DPRRHTVKAEBIGMISPLSLT-----PLVPMGRFTYA 347
DB 241 AFGCEILKYGTAWALREBELAAQKITMTLDEVREBRMNVLSITDTRPPIKGNPAVIAFV 300
QY 348 GIADLVHPRBOYTLMEHMGKPELIVWYFGGHTGF 383
DB 301 AATDGYIPKHSVLELQKAMPQSEVRAWYTGHSVF 336

RESULT 4
AAG20853
ID AAG20853 standard; protein; 356 AA.
XX
AC AAG20853;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SRQ ID NO: 23197.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hydridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 01-JUN-1999; 99US-0137528P.
PR 03-JUN-1999; 99US-0137752P.
PR 04-JUN-1999; 99US-0137752P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.

QY 251 RLSSIRSOEBSLIGLNGLSIGYIASIVASLEBGLACAIIGVPADLIIEILGRH---- 306
 Db 193 SLIHLMDTEEGFGKGVCSLWNGVNASVWGSILHP-----TVATL-PELSHSNAV 243
 QY 307 -C-----GLRHK-DPRRHTVMAEPIGSMISPLSLT-----PLVMPGRFIYA 347
 Db 244 AFCEGILKKYGTWMLREELAAQKITMTIDEVERRRNVLSLTDVTRFPINONPAVILFV 303
 QY 348 GIADRLVHPREQVTRLMEHMGKPEIWMYPGHTGPF 383
 Db 304 AATDGYIPKHSVLELQKAMPSESVRWVTGHSVSF 339

RESULT 5
 ADX77895
 ID ADX77895 standard; protein; 369 AA.

AC ADX77895;
 DT 21-APR-2005 (first entry)
 DE Plant full length insert polypeptide seqid 47261.
 XX
 XX plant protectant; plant growth regulant; gene therapy; plant;
 KM recombinant DNA construct; physical array; plant breeding marker;
 KM cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
 KM extreme osmotic condition; pathogen tolerance; pest tolerance;
 KM growth rate; cell cycle pathway; disease resistance;
 KM galactomannan production; lignin production; plant growth regulator;
 KM yield; plant growth; plant development; seed oil; protein yield;
 KM protein content.
 XX
 XX Unidentified.
 OS
 XX US2004034888-A1.
 PN
 XX 19-FEB-2004.
 PD
 XX 28-APR-2003; 2003US-00425114.
 PP
 XX 06-MAY-1999; 99US-00304517.
 PR 05-NOV-2001; 2001US-00985678.
 XX
 XX (LIUJ/) LIU J.
 PA (ZHOU/) ZHOU Y.
 PA (KOVA/) KOVALIC D K.
 PA (SCRE/) SCREEN S E.
 PA (TABAR/) TABASKA J E.
 PA (CAOY/) CAO Y.
 XX
 XX Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
 PI WPI; 2004-180133/17.
 DR
 XX New recombinant DNA construct, useful for improving plant tolerance to
 PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
 PT pests, for conferring increased resistance to plant disease, or for
 PT improving yield.
 XX
 XX Claim 1; SEQ ID NO 47261; 15pp; English.
 PS
 XX The invention describes a recombinant DNA construct comprising a
 CC polynucleotide consisting of a sequence encoding an amino acid sequence
 CC available in electronic form from the US patent office at
 CC ftp.segdata.uspto.gov/sequence.html?DocID:2004034888. The polynucleotide
 CC of the invention are also useful in physical arrays of molecules and as
 CC plant breeding markers. The recombinant DNA construct is useful for
 CC improving plant tolerance to cold, heat, drought, herbicides, extreme
 CC osmotic conditions, pathogens or pests, for manipulating growth rate in
 CC plant cells by modification of the cell cycle pathway, for conferring
 CC increased resistance to plant disease, for producing galactomannan,
 CC lignin or plant growth regulators, for increasing the rate of homologous
 CC recombination in plants, for improving yield by modification of

CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
 CC or by providing improved plant growth and development under at least one
 CC stress condition or for modifying seed oil or protein yield and/or
 CC content. This is the amino acid sequence of a plant full length insert
 CC polypeptide that can be used in the recombinant DNA construct of the
 CC invention.
 XX
 XX SQ Sequence 369 AA;
 Query Match 6.7%; Score 146.5; DB 8; Length 369;
 Beet Local Similarity 24.6%; Pred. No. 4.8e-06;
 Matches 59; Conservative 31; Mismatches 81; Indels 69; Gaps 8;

QY 208 NIWPEVL--PMHGPRGGLPRGA-----VFPGEVDLVDHGTAAVMDIRILSIR 257
 Db 149 NIATWLESPPYGGRRPMSRGAKLQCVSDLLLGKATIDBA-----RSLIWMIQ 198
 QY 258 SOEBSLIGLNGLSIGYIASIVASLE-----EGL----- 287
 Db 199 NEAGYKMGICGLSNGVHAMVGSILHPTPVATLPELAPHSNAVPPCEGVYKATAMDAL 258
 QY 288 --ACAILGVPADLIIEILGRHCGLRHKDPRRHTVMAEPIGSMISPLSLT-----PLVPM 340
 Db 259 KQDAAVLTIQDVTLLAEDAAQSGITIEQVR-----DRLSVLSLTDVTRFPVPRKN 308
 QY 341 PGRPIYAGIADRLVHPREQVTRLMEHMGKPEIWMYPGHTG--PFQSRPVRPFVQALAEQ 398
 Db 309 PQAVIFGATDGYITPRHSVWELQKAMPSESVRWVTGHSVSFFLHNDAPKRAIVDALDR 368

RESULT 6
 ADX68280
 ID ADX68280 standard; protein; 369 AA.
 XX
 XX ADX68280;
 DT 21-APR-2005 (first entry)
 DE Plant full length insert polypeptide seqid 39123.
 XX
 XX plant protectant; plant growth regulant; gene therapy; plant;
 KM recombinant DNA construct; physical array; plant breeding marker;
 KM cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
 KM extreme osmotic condition; pathogen tolerance; pest tolerance;
 KM growth rate; cell cycle pathway; disease resistance;
 KM galactomannan production; lignin production; plant growth regulator;
 KM yield; plant growth; plant development; seed oil; protein yield;
 KM protein content.
 XX
 XX Unidentified.
 OS
 XX US2004034888-A1.
 PN
 XX 19-FEB-2004.
 PD
 XX 28-APR-2003; 2003US-00425114.
 PP
 XX 06-MAY-1999; 99US-00304517.
 PR 05-NOV-2001; 2001US-00985678.
 XX
 XX (LIUJ/) LIU J.
 PA (ZHOU/) ZHOU Y.
 PA (KOVA/) KOVALIC D K.
 PA (SCRE/) SCREEN S E.
 PA (TABAR/) TABASKA J E.
 PA (CAOY/) CAO Y.
 XX
 XX Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
 PI WPI; 2004-180133/17.
 DR
 XX New recombinant DNA construct, useful for improving plant tolerance to
 PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or

XX PS Claim 2; SEQ ID NO 9602; 14pp; English.
XX
CC The invention relates a recombinant DNA construct comprising a
CC polynucleotide having any of 5544 nucleotide sequences (cdnas SEQ ID NO:
CC 1-5544) and encoding a polypeptide with any of 5544 amino acid sequences
CC (SEQ ID NO: 5545-11088). The cDNAs and proteins are from corn, soybean,
CC Arabidopsis, wheat and rape but the specification does not indicate which
CC sequences is derived from which organism. Also included is a method of
CC producing a plant having an improved property, comprising transforming a
CC plant with a recombinant DNA construct comprising a promoter region
CC functional in a plant cell operably joined to a polynucleotide encoding a
CC polypeptide associated with the property, and growing the transformed
CC plant. The property is selected from improving plant cold tolerance, for
CC manipulating growth rate in plant cells by modification of the cell cycle
CC pathway, for improving plant drought tolerance, for providing increased
CC resistance to plant disease, for galactomanan production, for production
CC of plant growth regulators, for improving plant heat tolerance, for
CC improving plant tolerance to herbicides, for increasing the rate of
CC homologous recombination in plants, for lignin production, for improving
CC plant tolerance to extreme osmotic conditions, for improving plant
CC tolerance to pathogens or pests, for yield improvement by modification of
CC photosynthesis, for modifying seed oil yield and/or content, for
CC modifying seed protein yield and/or content, for yield improvement by
CC modification of carbohydrate, nitrogen or phosphorus use and/or uptake
CC and for yield improvement by providing improved plant growth and
CC development under at least one stress condition. The polynucleotide may
CC also encode a plant transcription factor. The methods and compositions of
CC the present invention are useful in the field of biochemistry and
CC genetics, in particular for producing transgenic plants with improved
CC biological characteristics such as increased yield, improved nitrogen
CC flow, increasing plant tolerance to cold or heat, improving plant
CC tolerance to extreme osmotic and drought conditions, and improving plant
CC tolerance to plant pests or pathogens. They can also be used in physical
CC arrays of molecules, plant breeding markers, computer-based storage and
CC analysis systems. The present sequence is one of the 5544 plant protein
CC sequences of the invention. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?docID=20040216190.
XX
SQ Sequence 322 AA;
Query Match 5.4%; Score 118.5; DB 8; Length 322;
Best Local Similarity 25.3%; Pred. No. 0.0034;
Matches 57; Conservative 25; Mismatches 76; Indels 67; Gaps 10;
QY 208 NIWNPVL--PMHGRGGLPRKGA-----VPGEDVLDDVHGAQVWMIIRLLSIR 257
DB 130 NIATVLESPPYGGRRRPMQGAKECVSDLLLGKATIDDA-----RSLTWLQ 179
QY 258 SQEESLISGLNGLSGYIASLVEGLACAILGV-----VADLIELGRHC--- 307
DB 180 AEPYGMKGICGLMGVGHAAAGSL---LSTPIPLPLPHSAVVPFCGIVRHATW 236
QY 308 -GLRH-----KDRPRHTVKA-----EPGRMISPLSLPLVPMGRFVYAGIADRLVH 355
DB 237 EALREDAALAKDATSLEDAASGITTEQVXROIT-----DDGYI 276
QY 356 PREQVTLMEHMGKEPIVWPGCHTGP--QSRPVRFPVQAALPQ 398
DB 277 PRHSVMELOKAMPSEVWRVWTGVHSSFLHLNDSFRKAIYDALDR 321

RESULT 11
AB060724
ID AB060724 standard; protein, 889 AA.
XX
XX AB060724;
XX AC
XX DT 29-JUL-2004 (first entry)
XX
DE Klebsiella pneumoniae polypeptide seqid 7241.

XX KM Recombinant expression vector; transcription regulatory element;
XX KM Klebsiella pneumoniae protein; antibacterial; Vaccine.
XX OS Klebsiella pneumoniae.
XX PN US6610836-B1.
XX PD 26-AUG-2003.
XX PF 27-JAN-2000; 2000US-00489039.
XX PR 29-JAN-1999; 99US-0117747P.
XX PA (GENO-) GENOME THERAPEUTICS CORP.
XX PI Breton GL, Osborne M;
XX WPI; 2003-895346/82.
XX DR N-PSDB; ACH94275.
XX PT New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for
XX preparing a vaccine composition against Klebsiella pneumoniae.
XX PS Disclosure; SEQ ID NO 7241; 932pp; English.
XX
CC The invention describes a new isolated nucleic acid encoding a Klebsiella
CC pneumoniae polypeptide. Also described are: a recombinant expression
CC vector comprising the nucleic acid, operably linked to a transcription
CC regulatory element; and a cell comprising the recombinant expression
CC vector. The nucleic acid is useful for preparing a vaccine composition
CC against Klebsiella pneumoniae. This is the amino acid sequence of a
CC Klebsiella pneumoniae polypeptide of the invention
XX
SQ Sequence 889 AA;
Query Match 5.4%; Score 117.5; DB 7; Length 889;
Best Local Similarity 22.8%; Pred. No. 0.018;
Matches 100; Conservative 59; Mismatches 158; Indels 121; Gaps 26;
QY 20 PVLGAMG-----PMHTGGLVVAQSWRD--YLGOQDPKLPARPTTLAAQAFDEIVLL 72
DB 240 PVITALLNMGFIKPVTRQGL-----SPADVFPDPDSKK-----REASRFIDQVGP 289
QY 73 GL-----KARPVSNHNVFERISQEVAAGLEFYGNRRMLEKPSGFFAOPPELVAV 124
DB 290 GLVYLAEQMGKTEALVAAVQMLVQEKATGI-----YFALPTQLTS--- 332
QY 125 RKVYDRRSFYRTFFDSGFTPHGEPGSRMLSTANNREYALLRHREPRMLVCHGT 184
DB 333 NKTYDRNAPFLHOIV-STETP-----CHSLLL--HSGAWLM--DT 367
QY 185 EMGRAPLDLAVFRAMKHLDELGL-----NIWNPVLPM-HG-PRGGLPRGAVPFG 232
DB 368 EMGE---EGSPGAGPMHRRKGLAPRAVGTIOALMAVNVHGFRAAGLAGKV--- 421
QY 233 EDVLDDVH-----GT-AQAVMDIR-----LLSIRSQ-BEESLISGLNGSLGYIA 277
DB 422 --ILDEVHTYDLYGTITNALVFLRQIDCTVILISATLSQTRDALLOOSTSEAVPLI 479
QY 278 SLVASLEGLACAILGVPAVDLIELGRHCGLNHKDPRRTVMAEDIGMISPLSLT-- 335
DB 480 TAAPSAREBERGLVEIGVTEENTVTILHSC--RDEPARREBALRRRELGGQVLMIENTIA 537
QY 336 ----PLVPMGRFVIAGIADRLVHPR---EQVTLMEHMGKEPIVWPGCHTGFQSPRPV 388
DB 538 EAQOTYIDLASRAVACGTEGLHRSRTPOHRRHQRW-----VALY--GRAGHPQKQC 591
QY 389 RRFV--QAALQSGGLDA 404
DB 592 GRILVGTQVLEQSLDIDA 609

RESULT 12
 ID ABO71490 standard; protein; 409 AA.
 AC ABO71490;
 XX 29-JUL-2004 (first entry)
 DT
 DE Pseudomonas aeruginosa polypeptide #3665.
 XX
 KW Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
 XX
 OS Pseudomonas aeruginosa.
 XX
 PN US6551795-B1.
 XX
 PD 22-APR-2003.
 XX
 PF 18-FEB-1999; 99US-00252991.
 XX
 PR 18-FEB-1998; 98US-0074788P.
 XX
 PR 27-JUL-1998; 98US-0094190P.
 XX
 PA (GENO-) GENOME THERAPEUTICS CORP.
 XX
 PI Rubenfield MJ, Nolling J, DeLoughery C, Bush D;
 XX
 DR MPI: 2003-615309/58.
 DR N-PsDB; ABD05061.
 XX
 PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
 PT useful as molecular targets for diagnostics, prophylaxis and treatment of
 PT pathological conditions resulting from bacterial infection.
 XX
 PS Disclosure; SEQ ID NO 20236; 455bp; English.
 XX
 CC The invention relates to Pseudomonas aeruginosa polypeptides and the
 CC polynucleotides encoding them. The sequences are useful in diagnosis and
 CC therapy of pathological conditions, as molecular targets for diagnostics,
 CC prophylaxis and treatment of pathological conditions resulting from a
 CC bacterial infection, for evaluating a compound, such as a polypeptide,
 CC for the ability to bind a P. aeruginosa nucleic acid, as components of
 CC effective antibacterial targets, as targets for antibacterial drugs,
 CC including anti-P. aeruginosa drugs, as templates for recombinant
 CC production of P. aeruginosa-derived peptides or polypeptides, as target
 CC components for diagnosis and/or treatment of P. aeruginosa-caused
 CC infection, and in detection of P. aeruginosa sequences or other sequences
 CC of Pseudomonas species using biotech technology. Sequences ABO67826-
 CC ABO8436 represent P. aeruginosa polypeptides of the invention. Note: The
 CC sequence data for this patent did not form part of the printed
 CC specification but was obtained in electronic format from USPTO at
 CC seqdata.uspto.gov/sequence.html
 CC
 CQ Sequence 409 AA;
 XX
 XX

	Query Match	5.1%	Score 112.5	DB 7	Length 409;
	Beech Local Similarity	24.7%;	Pred. No. 0.02;		
	Matches	94;	Conservative	33;	Mismatches 111; Indels 143; Gaps 23;
OY		105	WLEKPSGFAOPPLTEVAVRKVCKRRRSFYRIPEDSG---- <td>158</td> <td></td>	158	
	:	:	:	:	:
Db		66	MDEVATGVIRRHHPAT-----ADNRR--RLFLALPGPRATLLEGHPG-GGSVPWGEVAT	114	
	:	:	:	:	:
OY		159	-----TNNREYALLLRHPPD-----RPMUCV-----	181	
	:	:	:	:	:
Db		115	GVIIRYPRAATADNRGRGLPALQAHVCRCRKATKGMFPFCMAITMTAGARCLCPAVILHRHNPQV	174	
	:	:	:	:	:
OY		182	----HGTEMGRAPLDLAYFRWKCHDELGLNTVM--PYL-----PMGGPR-----GQ	222	
	:	:	:	:	:
Db		175	DTDHRRVVVG--GPVRLGVVQAVLLFRGLDDVVVPHGLVGVDVYLHRAAYLARQAGVGE	233	
	:	:	:	:	:
OY		223	GLPKGAVFPEGEDV-----LDDVH-----GTAAVDMDIR--RLLSWIR	257	
	:	:	:	:	:

Db 224 GVVKKAGLPGEHVGTALSCARLTRNQGVGAGDAVDHVRFLGIGLEVADDAQLRLGGAR 293

Qy 258 SOEEBSLIGLNGLSLGGYIASLVASLEBGLACAILGVPADLLIBLGRHCGR--HKDP 315

Db 254 IARGVHHGLGSGARG-----QVAVTALAVAGVRA-----VAGRALRLQGVVRRDQ 339

Qy 316 RH-TVKMAEPIGRMISPLSLPLVMPGGRFYIAGIADRLVHREOVTRLMEH---WGKP 370

Db 340 AFAADLLLEGLGQ-----GRTV-AGVVELRVH-----RVEHRRGSIGAD 378

Qy 371 EIVWYPGHGTGF--QSRPVR 389

Db 379 LVAAIDQPHGNVAAQLRPVR 399

RESULT 13
 ABU34102
 ID ABU34102 standard; protein; 574 AA.
 AC ABU34102;
 DT 19-JUN-2003 (first entry)
 DE Protein encoded by Prokaryotic essential gene #19629.
 KM Antisense; prokaryotic essential gene; cell proliferation; drug design.
 OS Mycobacterium avium.
 PN WO200271183-A2.
 PD 03-OCT-2002.
 PF 21-MAR-2002; 2002WO-US009107.
 PR 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 PA (ELIT-) ELITRA PHARM INC.
 PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zvekind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 DR MPI; 2003-029926/02.
 DR N-PSDB; ACA37972.
 PT New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 PS Claim 25; SEQ ID NO 62026; 1766pp; English.
 XX
 XX
 CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene

CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 574 AA;

Query Match 5.1%; Score 112; DB 6; Length 574;
Best Local Similarity 22.5%; Pred. No. 0.036; Mismatches 148; Indels 168; Gaps 24;
Matches 105; Conservative 46; Mismatches 148; Indels 168; Gaps 24;

QY 23 SGAMGPFMTGLY-----VAQSMRDYLGQOPDKLPAPRTIALAAQFDEIYLGLKA 76
DQ 13 AGPLRPFADAGIFEAADIHVAQRLTALTGESDDRVALA---VALLVRLRGSGVCDLRA 69
QY 77 RRPVSNRVRERISQEV-----AAG-----LEFYG-----NRRL 106
DQ 70 -----VPAQVGAADLPMPAAGDWLAAYASPLGPPVLRFFGDLVYFDRYWL 117
QY 107 EK-----PSGFFAQ-----PPPLTEVAVRKVKDRRSFYRIFPDSGFTPH 147
DQ 118 EEEVCTDLALSPAGVSSCYERLEFPGEY-----QRAARIVSQALTTLT 168
QY 148 GEPSSQRLWSTYANNREYALL---RHPEPRMLVCVHGTGKRA----- 189
DQ 169 GGPETGK---TTTVARLLALLVEQERAGEBRPXIALAPGKAARLAABAIEHLD 225
QY 190 PLD---LAVPAMKLHDELGINIWPVLPMEGPRGQGLPKKAVPGEVDVDHGTAAV 246
DQ 226 PADARLRLAGLTGTLRLHGRPRPTSVAFKHN-RGNRLPHDI---VDETSMVSLTM 279
QY 247 WDIRRLSWIRSOEESLIGNGSLGVIASLVASLEGLACAILGVVADLIIEELGRH 306
DQ 280 --MARLAEAVRPDRLILVG-----DDQLASVAGAX-----LADLVGDLAGR 321
QY 307 CGLRHKPRRHTVKAER--IGRMISPLSLPLVPMGRFTYAGIADRLVHPRQVTRLM 364
DQ 322 AGVR-----VALATPHRFGSAIGALAAA-----IRAXDADRIV----- 355
QY 365 EHMGRPELVTPGSGHTGFPQS-RPVRRVQA-----ALESGILDA 404
DQ 356 -----ELLAGGHEIEMWDSERPADRLREVLVSHALRLRSALLGA 396

RESULT 14

ABO78103 standard; protein; 1127 AA.

XX ABO78103;

DT 29-JUL-2004 (first entry)

DQ Pseudomonas aeruginosa polypeptide #10278.

KM Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.

XX Pseudomonas aeruginosa.

XX US6551795-B1.

XX 22-APR-2003.

PF 18-FEB-1999; 99US-00252991.

XX

PR 18-FEB-1998; 98US-0074788P.
PR 27-JUL-1998; 98US-0094180P.
PA (GENO-) GENOME THERAPEUTICS CORP.
PI Rubentfield MJ, Nolling J, Deloughery C, Bush D;
XX WPI, 2003-615309/58.
DR N-PSDB; ABD11674.

PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.

XX Disclosure; SEQ ID NO 26849; 455bp; English.

CC The invention relates to Pseudomonas aeruginosa polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnosis,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC for the ability to bind a *P. aeruginosa* nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-*P. aeruginosa* drugs, as templates for recombinant
CC production of *P. aeruginosa*-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of *P. aeruginosa*-caused
CC infection, and in detection of *P. aeruginosa* sequences or other sequences
CC of Pseudomonas species using biolchip technology. Sequences ABO67826-
CC ABO84396 represent *P. aeruginosa* polypeptides of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html

SQ Sequence 1127 AA;

Query Match 5.0%; Score 109.5; DB 7; Length 1127;
Best Local Similarity 23.5%; Pred. No. 0.17; Mismatches 100; Indels 137; Gaps 18;
Matches 84; Conservative 36; Mismatches 100; Indels 137; Gaps 18;

QY 123 AVRKVDRRSFYRIFPDSGFTPHPGESQRLWSTYANNREYALLRHPEPR---WL 179
DQ 27 AARSIPERR-----PGRPCNR-----CTWTRSVSPLPFGSAPFASAW-- 65
QY 180 CNHGTWGRAPDLAVFRANKLH-DEGLNITWVLPMEGPRGQGLPKKAV-----FPG 232
DQ 66 ----SREGRCPCPYSCPTSRLEFVVDQERLELRIGIPVAHEGGQIGQEVVAHAGRAPV 121
QY 233 EDVLVDVHGTAAQVMDIRRLSWIRSOEESL-----IGNGSLGVIASLVASLEGLA 288
DQ 122 DRDADVHGLA-----VHLQPADALGHHRRRLDVAVGGDL-DLVA---GID 164
QY 289 CALGVVADLIIEELGRHCGLRHKDPRRHTVKAERPIGRMISPLSLPLV-----PMPGR 343
DQ 165 AQLLGRLADLDDELIG-----IGDRIQPAVLGVPEVGEPRVSGR 204
QY 344 FI-----YAGIADRL-----VHPRQVTL-----WENMGKPEI 372
DQ 205 RVGELVGRAEQFVULVLENPCRRVADRLAVVAVOGVHPRRLERLVVGERAFGLVDGEQ 264
QY 373 VWPVPG-----HTGPFQSRPVRRFVQAL-----EOSGLDAPRTQPD 410
DQ 265 ASHPFGVHDERVHGRF-----RRYVGLVGVGAAPGLAVPPHALLLVPRILAD 315

RESULT 15

ABG61612 standard; protein; 691 AA.

XX ABG61612;

DT 12-AUG-2002 (first entry)

DQ Human DRRP-3 splice variant #2.

DQ

XX Human; serine protease; dipeptidyl peptidase IV-related protein; DPP; DPP-IV; infection; human immunodeficiency virus; HIV-1; HIV-2; pain; diabetes; infertility; obesity; anorexia; Parkinson's disease; stroke; heart failure; hypertension; urinary retention; osteoporosis; cancer; ulcer; allergy; cancer; psychotic disorder; neurological disorder; dyskinesia; reproductive disorder; inflammatory disorder; metabolic disorder.

XX Homo sapiens.

OS WO200231134-A2.

XX 18-APR-2002.

XX 12-OCT-2001; 2001WO-US031874.

XX 12-OCT-2000; 2000US-0240117P.

XX (PERR) PERRING BV.

XX Qi S, Akinsanya KO, Riviere PJ, Junien J;

XX WPI; 2002-444178/47.

XX N-PSDB; ABK83343.

XX New dipeptidyl peptidase IV-related proteins and nucleic acids encoding the proteins, useful for treating e.g. fungal, bacterial, protozoan and viral infections, cancers, allergies, neurological disorders, or pain.

XX Disclosure; Page 107-109; 113pp; English.

XX The present invention relates to the isolation of novel human serine proteases referred to as dipeptidyl peptidase IV (DPP-IV)-related proteins (DPP-IV). The dipeptidyl peptidase IV-related proteins (DPP-IV) and nucleic acids encoding them are useful for treating infections such as fungal, bacterial, protozoan and viral infections, particularly infections caused by human immunodeficiency virus (HIV-1 or HIV-2), pain, diabetes, precocious puberty, infertility, obesity, anorexia, bulimia, Parkinson's disease, acute heart failure, hypotension, hypertension, urinary retention, osteoporosis, angina pectoris, stroke, ulcers, asthma, allergies, cancers, migraine, vomiting, psychotic, and neurological disorders (e.g. anxiety, dementia, or schizophrenia), and dyskinesias. These may also be used in discovering therapeutic agents for the treatment of reproductive, inflammatory and metabolic disorders. ABG61591-ABG61612 represent human DPP proteins

XX Sequence 691 AA;

XX

Query Match 4.9%; Score 106.5; DB 5; Length 691;
Best Local Similarity 18.2%; Pred. No. 0.18;
Matches 63; Conservative 63; Mismatches 106; Indels 115; Gaps 14;

Qy 9 THERSAPRLSPVLSGANGPMTHTGLVYAGSWRDYLGQOPDKLPYARPTIALAAQAFRDE 68
Db 288 SRDSKFFMTVIVKGGGGEFHIMFLIQSKSEQI-----TVRHILTSGNMEV 335
Qy 69 IVLGLKARRPVSNRVERISQEVAAGLEFYGNRMLEKPSGFPAPQPPLETVAVRKVK 128
Db 336 IKIL-----AYDETTOKISASTEGLNRQC-----ISCNPMK 367
Qy 129 DRRRSFYRIFPDGFTPHGEPGSGQRMVLSY-----TANREYAL----- 168
Db 368 EQ-----CTYFDASFSM-----NQHFLFCGGRPVVVSLSHTNPAKYFLBSNSMLK 417
Qy 169 ---LHHPBPRLVLCVHGTENGRAPLDLAVPRAWKIDELGINTVPLPMHGPRGGLP 225
Db 418 EAILKKKIGKPKIKILHIDY-ELPLQLSLPKDFMDRNOVALLLIMD-----EHPGQLVT 472
Qy 226 -----KGAVPGEEDVDDVH-----GTAQAVWDIRRL-----L 253
Db 473 DKFHIDMSVLIDMDNVIIVARPDGSGFGQIKIQEIHRRIGSVVEVDQITAVVFLKL 532

Qy 254 SWIRSQEBSLIGLNGLSGTYIASLVASLEGL-ACAILGVPAVL 299
Db 533 PYIDSKR-----LSIFKGYGGYIASMLKSDKLPKCGSVVAPITDL 575

Search completed: March 23, 2006, 05:10:56
Job time : 111.803 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 23, 2006, 05:11:18 ; Search time 16.4313 Seconds
(without alignments)
2418.401 Million cell updates/sec

Title: US-10-617-038-29

Perfect score: 2195
Sequence: 1 MASSASDGTHERSAPRLSP.....AALBOSGLDAPRTORDRSA 413

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: p1r:0:*
2: p1r1:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2195	100.0	413	2 B70573	hypothetical prote
2	293.5	13.4	400	2 C70837	hypothetical prote
3	126	5.7	582	2 D72536	probable acylamino
4	114	5.2	741	2 UC5142	X-Pro dipeptidyl-p
5	109	5.0	314	2 E72287	hypothetical prote
6	106.5	4.9	389	2 T30216	hypothetical prote
7	105	4.8	330	2 D97235	probable hydrolase
8	104	4.7	222	2 A82633	conserved hypotet
9	101.5	4.6	285	2 H72300	hypothetical prote
10	101.5	4.6	467	2 C72601	probable tryptopha
11	99.5	4.5	1303	2 C87519	hypothetical prote
12	97.5	4.4	318	2 B90458	hypothetical prote
13	97	4.4	1020	2 AB0701	tetrahionate redu
14	96	4.4	505	2 F71404	hypothetical prote
15	95	4.3	260	2 H84060	lysophospholipase
16	95	4.3	305	2 F69966	hypothetical prote
17	95	4.3	890	2 F75289	alanyl-tRNA synth
18	94	4.3	475	2 E90318	medium-chain-fatty
19	94	4.3	2082	2 T37056	probable multi-dom
20	93.5	4.3	359	2 B64919	conserved hypotet
21	93.5	4.3	359	2 D90820	hypothetical prote
22	93.5	4.3	359	2 A85769	hypothetical prote
23	93.5	4.3	1286	1 RJ0BP	interphotoreceptor
24	93.5	4.3	1668	2 A60272	IGA-specific metal
25	93.5	4.3	2324	1 A29924	acetyl-CoA carboxy
26	92.5	4.2	478	2 B72543	hypothetical prote
27	92.5	4.2	1212	2 F83153	probable two-compo
28	92	4.2	361	2 H75260	conserved hypotet
29	91.5	4.2	391	2 A97633	hypothetical prote

30	91.5	4.2	391	2 AD2856	conserved hypotet
31	91.5	4.2	1660	2 T17485	peptide synthetase
32	91	4.1	421	2 TC7367	second peroxisomal
33	91	4.1	633	2 T17262	hypothetical prote
34	91	4.1	1159	2 E83237	probable two-compo
35	90.5	4.1	307	2 C85363	hypothetical prote
36	90.5	4.1	502	2 S30916	bchx protein - Rho
37	90.5	4.1	502	2 T50754	prochlorophyllid
38	90.5	4.1	505	2 T03394	probable betaine-a
39	90.5	4.1	925	2 S27920	nuclear antigen EB
40	90	4.1	671	2 C81033	prolyl oligopeptid
41	89.5	4.1	400	2 H82833	proline dipeptid
42	89.5	4.1	570	2 D72597	hypothetical prote
43	89.5	4.1	593	1 A69655	two-component sens
44	89	4.1	357	2 C72746	probable carbamoyl
45	89	4.1	606	2 T35378	probable peptide h

ALIGNMENTS

RESULT 1

B70573 hypothetical protein Rv2627c - Mycobacterium tuberculosis (strain H37RV)

C/Species: Mycobacterium tuberculosis

C/Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 03-Jul-2004

C/Accession: B70573

R/Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churruarin, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Felcner, S.; Hamlin, N.; Holtrold, S.

Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skellon, S.; Squares, S.

Nature 393, 537-544, 1998

A/Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A/Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A/Reference number: A70500; MUID:98295387; PMID:9634230

A/Accession: B70573

A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-413 <COL>

A/Cross-references: UNIPROT:006185; UNIPARC:UP100000CCAP2; GB:295387; GB:AL123456; NID:9

A/Experimental source: strain H37RV

C/Genetics:

A/Gene: Rv2627c

Query Match	Score	Match	Length	ID	Description
Best Local Similarity	100.0%	2195	413		
Matches	413	Conservative	0	Mismatches	0
				Indels	0
				Gaps	0
QY	1	MASSASDGTHERSAPRLSPVLISGAMGPPMATGLVVAQSWRDYLCQQDPKLPAPRTIAL	60		
DB	1	MASSASDGTHERSAPRLSPVLISGAMGPPMATGLVVAQSWRDYLCQQDPKLPAPRTIAL	60		
QY	61	MAQAFRDELIVLGLKARPRVSNHRVPERISQVVAAGLEFYGNRRMLEKSGSFAPPPPT	120		
DB	61	MAQAFRDELIVLGLKARPRVSNHRVPERISQVVAAGLEFYGNRRMLEKSGSFAPPPPT	120		
QY	121	EVAVRKVKDRRSFYRIFPDGFTPHGPGSQRMVSTYANRREVALLRHSEPPWLVLC	180		
DB	121	EVAVRKVKDRRSFYRIFPDGFTPHGPGSQRMVSTYANRREVALLRHSEPPWLVLC	180		
QY	181	VHGTMGRAPLDLAVFRANKLHDELGLNIVMVLPMHGERGGCLPKGAVFPEEDVLDDVH	240		
DB	181	VHGTMGRAPLDLAVFRANKLHDELGLNIVMVLPMHGERGGCLPKGAVFPEEDVLDDVH	240		
QY	241	GTAQAVMDIRRLLSVIRSQEESLSGLNGLSIGVYIASIVASLEGLACALLGVPAADI	300		
DB	241	GTAQAVMDIRRLLSVIRSQEESLSGLNGLSIGVYIASIVASLEGLACALLGVPAADI	300		
QY	301	ELLGRHCGLRHDPRRHTVKAEPYIGRMISPLSTPLVPMGRFTYAGIADRLVHPRQV	360		
DB	301	ELLGRHCGLRHDPRRHTVKAEPYIGRMISPLSTPLVPMGRFTYAGIADRLVHPRQV	360		
QY	361	TLVHMHGKPEIVWYPGGHTGFQSGRPVRRVQAALBOSGLDAPRTORDRSA	413		
DB	361	TLVHMHGKPEIVWYPGGHTGFQSGRPVRRVQAALBOSGLDAPRTORDRSA	413		

Db 361 TELMEHWKPEIWMVFGHTFFQSRPVRVQVALLGSGLLDAPRTQRDRSA 413

RESULT 2

C70837

hypothetical protein Rv0293c - Mycobacterium tuberculosis (strain H37Rv)

C:Species: Mycobacterium tuberculosis

C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004

C:Accession: C70837

R:Coile, S.T.; Brosch, R.; Parthill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltham, T.; Gentles, S.; Hamlin, N.; Holtrold, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skellern, S.; Squares, S. Nature 353, 537-544, 1998

A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A:Reference number: A70500; MUID:98295987; PMID:9634230

A:Accession: C70837

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-400 <COL>

A:Cross-references: UNIPROT:O53697; UNIPARC:UP100000C1548; GB:AL021930; GB:AL123456; NID

C:Experimental source: strain H37Rv

C:Genetics:

A:Gene: Rv0293c

Query Match 13.4%; Score 293.5; DB 2; Length 400;

Best Local Similarity 30.5%; Pred. No. 1e-16;

Matches 97; Conservative 42; Mismatches 122; Indels 57; Gaps 14;

Qy 95 AGLPEYGNRMLEKSGFPAQP-----PRLTEVAVKRVDRRSFRIIPDSGF 143

Db 84 SNLNFYAEALAAHDPKSPAPTELPKTSRPASPLTEWAVAGTVD-----NIAASGF 137

Qy 144 TP-HGCEPQSQWMLSTYANNREYALLLRHP-PRPLVCVHGTEMGRAPLDLAVPRAWL 201

Db 138 RAINEP-TMRQMSALTANNIYHAGWKRDRGPRFLCYIHGF-MQSSITLNGLPFSLEW 194

Qy 202 HDELGLTVMVPLPMHGPFGQGLPKGAVPFG-----EDVLVDVHGTAQVMDIRLLSMI 256

Db 195 YRSRGVDVLYTLPLFGQARF---KSPFSFGYFTSGLSGFAMAQAVVDYFRSIVDYL 251

Qy 257 RQSEBSLIGLNGLSIGYTIASIVASLEGLACALIGVVA-----DLIELL 303

Db 252 RHIGVDR-IALTGISIGYTSALLASVESRLVENCVVMPAKLFDEMPANKLVKL- 309

Qy 304 GRHCGLRHNDPRRHVTYKMAEPICGMI--SPLSLTPLVPMGPFYAGIADRLVHPREQYT 361

Db 310 ----GLRLTN-----ISRDELINGLAVHGPFLVWRPLLPDRKMITITGLDRAAPRHHAVT 360

Qy 362 RLMEHWKPEIWMVPGCH 379

Db 361 -LWKQWDRCALHWPQSH 377

RESULT 3

D72636

probable acylamino-acid-releasing enzyme APE1547 - Aeropyrum pernix (strain K1)

C:Species: Aeropyrum pernix

C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004

C:Accession: D72636

R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Hatake, Y.; Jin-no, K.; Takahara, H.; Takamiya, M.; Maeda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K DNA Res. 6, 83-101, 1999

A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum

A:Reference number: A72450; MUID:99310339; PMID:10382966

A:Accession: D72636

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-582 <KAW>

A:Cross-references: UNIPROT:Q9YBQ2; UNIPARC:UP1000005DFBF; DDBJ:AF000061; NID:G5104821;

C:Experimental source: strain K1

C:Genetics:

A:Gene: APE1547

Query Match 5.7%; Score 126; DB 2; Length 582;

Best Local Similarity 25.4%; Pred. No. 0.016;

Matches 86; Conservative 36; Mismatches 130; Indels 86; Gaps 20;

Qy 37 AQSMVDYLGQDP-KLPIARPITLAAAFDEIYLLGLKARRPVSNN-RVPERISQVYA 94

Db 247 ALTWLGYL---PDGLAV-----VAREGSNAVIDEREAAPQNGRVLVWRKLVLT 297

Qy 95 AGLPEYGNRMLEKSGFPAQPPLTEVAVKRVDRRSFRIIPDSGFTPHPGSGQR 154

Db 298 SHITSSTPRIVSLPSG-----EPLLEGL--PELRKSI-----AGRL 335

Qy 155 -WLSTYANNREYALL---RHPEPRMVLVCHGTEMGRAPLDLAVFRAMKLDLGLNIV 210

Db 336 VVVFESFDSRVPTVYVLESGRAPTPGPTVVLVHGFABDSMDTFASLA--AAGFHV 393

Qy 211 MPVLPMAHGPFGQL-----PKGAVFPREDVDVHGTAQVMDIRLLSMIRSGE 261

Db 394 MP--MYRSTGYGEBMRKIIIGDPCG---GE--LEDVSAAR-----MARESGL 435

Qy 262 ESLIGLNGLSIGYTIASIVASLEGL-ACALIGVPVAD-----LIELL--G 304

Db 436 ASELTYMGYSIGYITLCAITWKPGFLFAQVAGASVVDMEEMTELSDAAPNFIQLTGG 495

Qy 305 RHCGLRHNDPRRHVTYKMAEPICGMI--SPLTPLVPM 340

Db 496 SREIMRSRSPINHVDRIKEPLA-LIHPONDRSTPLPKL 532

RESULT 4

JC5142

X-pro dipeptidyl-peptidase (EC 3.4.14.11) - Xanthomonas maltophilia

C:Species: Xanthomonas maltophilia

C:Date: 31-Jan-1997 #sequence_revision 27-Feb-1997 #text_change 09-Jul-2004

C:Accession: JC5142

R:Kadashima, T.; Ito, K.; Yoshimoto, T.

A:Biochem. 120, 1111-1117, 1996

A:Title: Dipeptidyl peptidase IV from Xanthomonas maltophilia: Sequencing and expression

A:Reference number: JC5142; MUID:97164011; PMID:9010758

A:Accession: JC5142

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-741 <KAB>

A:Cross-references: UNIPROT:P95782; UNIPARC:UP100000B5F8A; DDBJ:D83263; NID:G1753196; PI

C:Comment: This enzyme catalyzes the hydrolytic removal of amino-terminal dipeptidyl res

C:Superfamily: dipeptidyl-peptidase IV

C:Keywords: dipeptidylpeptide hydrolase; membrane bound

F:4-18/Domain: transmembrane #status predicted <TM>

F:610/Active site: Ser #status predicted

F:685/Active site: Asp #status predicted

F:717/Active site: His #status predicted

Query Match 5.2%; Score 114; DB 2; Length 741;

Best Local Similarity 21.5%; Pred. No. 0.21;

Matches 102; Conservative 57; Mismatches 167; Indels 148; Gaps 25;

Qy 34 LYVAQ-SMRD-----YLQGPDP--KLPIARPITLAAQ-----APRD 67

Db 287 IYLAIVNKRDDQHLSPQSRDQKLDLVEYTIASNOQRYLAHETSPVWPLHNSLRLLD 346

Qy 68 EIVLD-----GLKARRPVSNHRFERISQEVAGLEFYGNRWLE--K 108

Db 347 DGSILMSERTGFQHLVIRIDSKGKAALLTHGWSVDELAVDEKAGLAVF--RAGIESAR 404

Qy 109 PSGFPA-----QPPPLTEVAVRKVD--RRRSFRIIPDSGFTPHPGSGQRWLSTYA 160

Db 405 ESQIYAVPLQGGQPPQLSKAPGMSAPFARNAVYVDSWNSSTPPIE-----LPPA 457

Qy 161 NNREVALLL-----RHPEPRMVLVCHGTEMGRAPLDLAVFRAMKLDLGLNIVMP- 212

Db 458 NGEKIATLVENDLADPKPRVYR-----YREARVPEFGTTLTAADGKTPLNYSVIRPA 509

QY 213 -----VLPNHGPRGQ-----GLP-KGAV 229
 DB 510 GEDPAKRYVAVAYVYGQVITQDTPGSGDHLFNQYLAQGVVPSLDNRGTSPRRGD 569
 QY 230 PGEVDVLDVHGTGAQAVMDIRRLLSWIRSGOE--EESLIGLNGSLGYSIA-SLVALEEG 286
 DB 570 FGG--ALYKQGVTE-VADQLKGVAMLKQQPMWDPRARIGVQSGNSGVTYMLLAASIS 626
 QY 287 LACAILGVPAVDLIELGRHCGLRHND-PRRHTVKAEP-----IGRMTSPSLTPLVPM 340
 DB 627 YACGVAGAVPTDW-GLYDSHYTERYMDLPRANDAGREARVLTHIGLSPL----- 677
 QY 341 PGRFYIAGIADR-LVHPREQVRLMEHWKRP-EIYVPGSGHTGFQSPRPVRRF 391
 DB 678 ---LTHGMADNVLTFTNSTLSMALQKGGQPEELMTYGAARKGLSGDALHRY 728

RESULT 5

E72287
 hypothetical protein - Thermotoga maritima (strain MSB8)
 C:Species: Thermotoga maritima
 C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
 C:Accession: E72287
 R.Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey, Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.; C.M.
 Nature 399, 323-329, 1999
 A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
 A:Reference number: A72200; M0ID:99287316; PMID:10360571
 A:Accession: E72287
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-314 <ARN>
 A:Cross-references: UNIPROT:Q9X0P3; UNIPARC:UP100000C1286; GB:AE001773; GB:AE000512; NID
 C:Genetic8:
 A:Experimental source: strain MSB8
 A:Gene: TM1159

Query Match 5.0%; Score 109; DB 2; Length 314;
 Best Local Similarity 23.7%; Pred. No. 0.19;
 Matches 61; Conservative 34; Mismatches 97; Indels 65; Gaps 14;
 QY 195 VERAWKLIH--DELGLNIWVPLPMHGPRGQGLPKGAVPGEVDVLDVHGTGAQAV----- 247
 DB 70 LYLIMAHYLSRKGIRAILPVLPGNPTR---VAEGSV-SGKDYFSSDLRMSRFWEHALT 125
 QY 248 DIRRLLSWIRSGOE--EESLIGLNGSLGYSIASLVALEEGACAILGVPAVDLIEL- 303
 DB 126 DLJLSLELLKVKMKMHERNC--LFGYCLGMIALVNLNALSDDPKTTIMAGDEPATLFW 183
 QY 304 -----GRHCGLRHK---DPRRHTVK-----MAEPIGRMTSP-TSLT 335
 DB 184 KAPTLSFVRRELKSGGEHGMENENFVDYRSDLERLKEPSSVQEMSSNHLPLKLD 243
 QY 336 PLVMPGREFIYAGIAD-RLVH-----PREQVRLMEHWKRPDIYVPGSGHTGF-- 382
 DB 244 PLA-----YAKFVDTSRIVMLEAWFDRALPKSTRDILMEHLGKFKRIKVPSSHSWLP 296
 QY 383 FOSRPVRFRVQALBQS 399
 DB 297 FQVLARLYIMKLKEQS 313

RESULT 6

T30216
 hypothetical protein Z - Streptomyces hygroscopicus
 C:Species: Streptomyces hygroscopicus
 C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 09-Jul-2004
 C:Accession: T30216
 R.Aparticio, J.F.; Molnar, I.; Schwecke, T.; Koenig, A.; Haydock, S.F.; Khaw, L.E.; Stuart
 Gene 169, 9-16, 1996
 A:Title: Organization of the biosynthetic gene cluster for rapamycin in Streptomyces hyg
 A:Reference number: Z20782; M0ID:96186896; PMID:8635756

A:Accession: T30216
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-389 <APA>
 A:Cross-references: UNIPROT:Q54287; UNIPARC:UP100000BELFA; EMBL:X86780; NID:9987088; PID
 C:Superfamily: Streptomyces hygroscopicus hypothetical protein Z

Query Match 4.9%; Score 106.5; DB 2; Length 389;
 Best Local Similarity 24.1%; Pred. No. 0.4;
 Matches 71; Conservative 25; Mismatches 122; Indels 77; Gaps 13;
 QY 113 FAQPPPLVEAVARKYK---DRRSFYRIFFPSGFTPHGSEPSQRMSTYANNEVALL 168
 DB 112 FADAAPLDLTPRAAIRPIEDTTLPGYLFVDDSGTTPR----- 150
 QY 169 LNRHPRPRLVCVHSTEMGRAPLDLAVFRAMKDELGLNIWVPLPMHGPRGQ---LP 225
 DB 151 -----TVLFHGQYSTLEEDYLAALAGL--RQGN-----VIAFDGP-GQGSYARE 194
 QY 226 KCAVPPGEVDVLDVHGTGAQAVMDIRRLLSWIRSGOEESLIGLNGSLGYSIASLVALEEG 285
 DB 195 QGHNFP-----PDWEAVVTVPVDFALTLPEV---DAEQLY-LMGKSLGYSIAARAAPFH 245
 QY 286 GLACAILGVPAVDLIELGRHCGLRHNDPRRHTVMAEPI-----GRMTSPSL-LTP 336
 DB 246 RVAAACVLVDGYDFHVAIVATAGRAASTPGGMEALMAQTVRWVYVNGRWTTFGVSQVDE 305
 QY 337 IVPMPGREFIYAGIADRLVHPREQVRLMEHWKRPDIYVPGSGHTGFQSPRPVRRF 391
 DB 306 LVRAAEPTTMAGIARITCP-----TLVLE-----AENDGFPGQPPRLF 345

RESULT 7

D97235
 probable hydrolase from alpha/beta family, YOKD B. subtilis ortholog [imported] - Clostr
 C:Species: Clostridium acetobutylicum
 C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 05-Oct-2004
 C:Accession: D97235
 R.Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, J.; Daly, M.D.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
 J. Bacteriol. 183, 4833-4838, 2001
 A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo
 A:Reference number: A96900; M0ID:21359325; PMID:21359325
 A:Accession: D97235
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-330 <KUR>
 A:Cross-references: UNIPROT:Q97FLL; UNIPARC:UP100000CA5FA; GB:AE001437; PIDN:AAK80671.1;
 A:Experimental source: Clostridium acetobutylicum ATCC824
 C:Genetic8:
 A:Gene: CAC2725
 C:Superfamily: Alpha/beta hydrolase

Query Match 4.8%; Score 105; DB 2; Length 330;
 Best Local Similarity 21.1%; Pred. No. 0.43;
 Matches 44; Conservative 37; Mismatches 82; Indels 46; Gaps 8;
 QY 173 EBRPWLVCVHSTEMGRAPLDLAVFRAMK-----LHDELGLNIWVPLPMHGPRGQGLPK 226
 DB 89 DSRKTIIVHGTGSR-----WESMKYADIIYLDGIVNVLVYDSRYHVGSG----- 133
 QY 227 GAVPGEVDVLDVHGTGAQAVMDIRRLLSWIRSGOEESLIGLNGSLGYSIASLVALEEG 285
 DB 134 -----GNDI---TLGYFER-YDLANNVVMKVKKTGPGIIGIHGESEMGATATLQSNMEX 184
 QY 286 --GLACAILGVPAVDLIELGRHCGLRHNDPRRHTVMAEPIGRM-----ISF 331
 DB 185 TKDVFYVVDGCFPSDLPLDFGEKLVNEMKNHGAVAVKAVVFYSLSIAEFKAGFSVYALSP 244
 QY 332 LSLTFLVMPGREFIYAGIADRLVHPREQV 360
 DB 245 IKAIDVDKTPIMFAH-GADDDLIPENSV 272

RESULT 8

A02633 conserved hypothetical protein XP_01929 [imported] - Xylella fastidiosa (strain 9asc)

C|Species: Xylella fastidiosa

C|Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 05-Oct-2004

C|Accession: A02633

R|Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequence
Nature 406, 151-157, 2000

A|Title: The genome sequence of the plant pathogen Xylella fastidiosa.

A|Reference number: A02515; MUID:20365717; PMID:10910347

A|Note: for a complete list of authors see reference number A59328 below

A|Accession: A02633

A|Status: preliminary

A|Molecule type: DNA

A|Residues: 1-522 <SIM>

A|Cross-references: UNIPROT:Q9PCF2; UNIPARC:UPI0000C284A; GB:AEO04004; GB:AEO03849; NITD
R|Experimental source: strain 9asc

B|Rimponen, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvaresga, R.;
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargos, L.E.A.; Carraro, D.M.; Carter, H.
as-Neto, E.; Docena, C.; El-Dorriy, H.; Facinanti, A.F.; Ferreira, A.J.S.

S|Submitted to GenBank, June 2000

A|Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm,
J.D.; Junqueira, M.L.; Kemper, E.L.; Ktajiama, J.P.; Krueger, J.B.; Kurumae, E.E.; Laigret,
Chado, M., Medeira, A.M.B.N.; Madeira, L.T.; Marques, M.V.; Martins, B.M.; Mayaki, C.Y.;
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.B.; de Sa, R.G.; Santelli, R.V.; Sawasaki,
M.; Tenuko, M.H.; Vallada, H.; Van Sluyts, M.A.; Verjovski-Almeida, S.; Vetore, A.L.; Z

A|Reference number: A59328

A|Contents: annotation

C|Genetics:

A|Gene: XF1929

C|Superfamily: uncharacterized conserved protein

Query Match 4.7% | Score 104; DB 2; Length 222;
Best local similarity 25.0%; Pred. No. 0.31;
Matches 38; Conservative 24; Mismatches 60; Indels 30; Gaps 8;

OY 228 AVFGEDVLDVDHGTGAQAVMDIRLLSWIRSQEBSLIGLGSLSGYTASLVASLEBGL 287
 ::::|||||||:~::~|||:
DB 81 SVGASDMGFNDGNHGERR--DLRAIAAWVRAPDPSPALMLAGFSYGAYISLVAAELEBPQ 137
OY 288 ACATIGVVADLIELGHCHGRKHDPRRRHYKAEPYGMWISP-----LSLP----L 337
 ::|||:|||||||:~::~|||:
DB 138 VLIISSPA-----GR-WPSSHVPHEPHMLLTGGDADEVDDPGVAIVDWISTTPROPKL 189
OY 338 VPMP--GRFIYAIGIADRILVHPRECV---TRLN 364
 :|||:|||:|||||||:~::~|||:
DB 190 IRMPDTSHFPFR---KIHLRDALQGDSRW 217

RESULT 9

H72300 hypotheical protein - Thermotoga maritima (strain MSB8)

C|Species: Thermotoga maritima

C|Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004

C|Accession: H72300

C|Netelson, K.E.; Clayton, R.A.; Gill, S.R.; Gunn, W.L.; Dodson, R.J.; Haft, D.H.; Hickey,
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;
C.M.

Nature 399, 323-329, 1999

A|Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome sequencing
of Thermotoga maritima.

A|Reference number: A72200; MUID:99287316; PMID:10360571

A|Accession: H72300

A|Status: Preliminary

A|Molecule type: DNA

A|Residues: 1-285 <ARN>

A|Cross-references: UNIPROT:Q9XOD6; UNIPARC:UPI0000CC1315; GB:AEO01765; GB:AEO00512; NITD
A|Experimental source: Strain MSB8

C|Genetics:

A|Gene: TM1046

```

Query Match 4.6%; Score 101.5; DB 2; Length 285;
Best Local Similarity 21.7%; Pred. No. 69;
Matches 65; Conservative 44; Mismatches 133; Indels 57; Gaps 15;

OY IFPDGFTPHGEFGSQRWLSTYTANNEYALLLHPPEPMLVCVHSTEMGRAPLDLVF 196
DB ITPDASVDPLAESEKTRVVVFEEPERNLIF-----LHGIGNGP-----Y 47
OY 137 RAM--KLHDEIGLNIVMPVLPMGPRG-QGLPKGVFPGEDVDLDVHTGAQAVDIRLL 253
DB 48 LLMGCKFRREYNITKWTEILIPYHEKRAPENWSGGEPPFYHSSPSRCVKRPFDAVDIDL 107
OY 254 SWINSQEBSJLGNGSLGGYIASLVAS----LEEGLAAILG-----VPVADIE 301
DB 108 DVVR-RDDDKPIISLMGSFGSMITTIALAEKRILEKVICCGSGDMRWIMYSBYTERLR 166
OY 302 LL-----GRHCGLR-HKD---PRRT-----VKNAEPFG-RMISPLSLTPLVM 340
DB 167 ELYKNKGNEVGCRSEKOCIKRRRAAPELIKSFNSEDIRKKRPVGCYFDPAFAFPVDQ 226
OY 341 PGRITVGIADRLVHPREOVTRLMEHWKKEIYVPCGHG--FFQSRRPVRFQAAL 397
DB 227 KLFEFWA-LFDHVI-PLYQYACFLKLIKNNKTIVLPBGHGSYFFRRYIAKRVRFID 283

RESULT 10
C72601
Probable cryptophanase ABE1275 - Aeropyrum pernix (strain KI)
C1species: Aeropyrum pernix
C1date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C1accession: C72601
R.Kawarabayasi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haitkawa, Y.; Jin-no, K.; Takahara, H.; Takamaya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kawa, R., 83-101, 1999
A.Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum
A.Reference number: A745450; MUID:99310339; PMID:10382966
A.Accession: C72601
A.Status: preliminary
A.Molecule type: DNA
A.Residues: 1-667 <KMW>
A.Cross-references: UNIPROT:Q9YC12; UNIPARC:UP1000005DEA6; DDBJ:AP000061; NID:g5104821,
A.Experimental source: strain KI
C.Genetics:
A.Gene: ABE1275
C.Superfamily: tryptophanase

Query Match 4.6%; Score 101.5; DB 2; Length 467;
Best Local Similarity 22.5%; Pred. No. 1.3;
Matches 75; Conservative 48; Mismatches 106; Indels 105; Gaps 19;

OY 65 FRDEI-VLLGKARKRPNRNHVPERI-----SQEVAAGLEFGNRMLEKPGSF- 112
DB 86 FRDVRDVLGLDLVLPVHQGRAERILYGELLRRNRRIVPANTHTFDTGRAVILNQGV 145
OY 113 PAOPPLETAVAKRKORRSRYRIFPDGFTPHGE---PGSQRWLSYTANNEYALL 169
DB 146 LDLESP-----QASREAY-----PPKGDIDVARLERILK--ERSRDVAFIL 185
OY 170 RHPEPRLVLCVHSTEMGRAPLDL-AVFRWKCLHDELGLNIWVPV-----LPMGRP 219
DB 186 -----LVTTNTTAGGQ-FVSMNDVKTAREILDADAGLPLVMDICFAENAVLYKERDP 236
OY 220 RGQGLPKGVNFPGEDVDLDVHGTGAQAVMDI-----RLLSW-----IRSQEBSJLGNGLS 271
DB 237 RYRG-----WSVPDIAREMTSYGDHFVSAAKQGL-----AN 268
OY 272 LGGYIASLVASLEGLAAIL-----GVPAVDIELIGHHCGLR--HKDPRRHVT 319
DB 269 IGGYIAIRDPSLYEDLAAARVYLBEQYTVTGGLAGRDL-EAIAQ--GLREVVEEDYLHRV 325
OY 320 KMAEPFGMISPLSLTPLVMGFRFYAGIADRLL 353

```

Db 326 EQVRYKGLLSQGVPIVEPVGGHAYVVDLEAL 359

RESULT 11

C87519 hypothetical protein CC2180 [imported] - Caulobacter crescentus

C.Species: Caulobacter crescentus
C.Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #ext_change 09-Jul-2004
C.Accession: C87519

R.Niemman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.B.; Eiken, J.; Heidelberg, J.B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwim, M.L.; Haft, D.H.; Koonin, J.; Esmailova, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M. Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A.Title: Complete Genome Sequence of Caulobacter crescentus.

A.Reference number: A87249; MUID:21173698; PMID:11259647

A.Accession: C87519

A.Status: preliminary

A.Molecule type: DNA

A.Residues: 1-1303 <STO>

A.Cross-references: UNIPROT:Q9A6B4; UNIPARC:UPI00000C762F; GB:AE005673; NID:g13423679; H

C.Genetics:

A.Gene: CC2180

Query Match 4.5%; Score 99.5; DB 2; Length 1303;
Best Local Similarity 21.2%; Pred. No. 7.2;
Matches 89; Conservative 54; Mismatches 138; Indels 139; Gaps 23;

QY 1 MASSADGTHERSAPFLSPVLSGAMPFMHGLVVAQ-----SMRDYLGQCPDKLP 52

DB 422 VASRAKD--QQRVASLDDPLKGGVL-----FTTVAQETPIQELWAYVFGAEPK--D 472

QY 53 IARPTALAAQAFRDSIVLLGLKARAPVSNHVRFERISQVAAGLFEGYGRW-LEKPEG 111

DB 473 ARKMSYTVABAVADLVNLADLN-----YIAGRPMKERRATVVALPBG 516

QY 112 PFAQPPLETVAVRKVKRRRSFYRIFFDSGFTPH-GEPSGRWLSTYANNREVALLR 170

DB 517 AENRPAPADGLAAKGL-----PIVHLIPSGFGDAPAAKPLARW----- 556

QY 171 HEPFRWLVCVGTENGRAPLDLAVFRANKHDELGLNIVMPRLPMHGRGGLP----- 225

DB 557 -----AYGWEVARQGLD-----GVALDIPLP--GAAGAVIPLNIQV 591

QY 226 KCAVFPGEVDLD---DVHGTAQAVW-DIR-RLSWIRSQEBSLIGLNGLSGGYIASL 279

DB 592 KQPIPGRMIDVSVSKAGEARTWLDLRDLIT-----ADSL-----YLITA 635

QY 280 VASLEBG--LACALIGVEVADLIELGRHCGLRHNDPR-----RRTVKAEPFGM 328

DB 636 ADSAERGPALDQAKRLVLKDRDQAKVHIQDRLNQVDMNAFLVEBHTTSKQGLYRR 695

QY 329 ISPLSLTLPVPMGRITVYGIADRL-VHREOVTRLMHWKPEIYWGCHTGFQSPR 387

DB 696 -----LVADISDLRLVDPDNKIGR--EYWG--DIYVSGQWPAFEPQPK 735

RESULT 12

B90458

hypothetical protein SSO2810 [imported] - Sulfolobus solfataricus

C.Species: Sulfolobus solfataricus

C.Date: 24-May-2001 #sequence_revision 24-May-2001 #ext_change 09-Jul-2004

C.Accession: B90458

R.She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Aways, M.J.; Chan-

Jong, I.; Ueffler, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thl-NGOC, H.P.; Redder, F.

arrett, R.A.; Ragan, M.A.; Sengen, C.W.; Van der Oost, J.

Submitted to Genbank, April 2001

A.Description: Sulfolobus solfataricus complete genome.

A.Reference number: A99139

A.Accession: B90458

A.Status: preliminary

A.Molecule type: DNA

A.Residues: 1-318 <KUR>

A.Cross-references: UNIPROT:Q97V26; UNIPARC:UPI0000064872; GB:AE006641; NID:g13816160; H

C.Genetics:
A.Gene: SSO2810

Query Match 4.4%; Score 97.5; DB 2; Length 318;
Best Local Similarity 21.4%; Pred. No. 1.7; Indels 79; Gaps 10;
Matches 54; Conservative 32; Mismatches 79; Indels 87; Gaps 10;

QY 185 ENKRAPLDL-AVFRANKLHDELG-----INTVPLVPMHGRPGQGLPKGAVFGEVDL 236

DB 103 EGYTAPIDLGQALIMWKVDEKVAAYKFGDLNVKVPVIFPFGVIGTGLP----- 151

QY 237 DVHGTAQAVWDIRRLSWINSQEBSLIGLNGLSGG--YIASVLSLEGLACILGV 294

DB 152 -----FRKSLSTIPPRENGGMMDKHLTVGKYLIPFVNG-----ALISI 192

QY 295 PVADLIELGRHCGLRHNDPRHNTKMAEPGRMISPLST-PLVPMGRFIYAGIADRL 353

DB 193 GDTHLAQDGEVCGTAIEPMEVTVK-----RLKNVGLTLPF----- 232

QY 354 VHPREOVTRLMHWKPEIYWGCHTGFQSG-----RPVRRPYQA-- 395

DB 233 -----VTKVKEMEYDEVIAVPGIDSNLMNATKKAIGIISLSKMAPEVAYILASV 286

QY 396 --LEQSGLIDAP 405

DB 287 VMLRVSEVDVP 298

RESULT 13

AE0701

tetrahionate reductase chain A [imported] - Salmonella enterica subsp. enterica serovar

C.Species: Salmonella enterica subsp. enterica serovar Typhi

A.Note: this species has also been called Salmonella typhi

C.Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #ext_change 18-Nov-2002

C.Accession: AE0701

R.Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,

th, T.; Conerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,

S.; Moulie, S.; O'Gaora, P.

Nature 413, 848-852, 2001

A.Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;

A.Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov

A.Reference number: AB0502; MUID:21534947; PMID:11677608

A.Accession: AE0701

A.Status: preliminary

A.Molecule type: DNA

A.Residues: 1-1020 <PAR>

A.Cross-references: UNIPARC:UPI0000059DB8; GB:AL513382; PIDN:CAD01982.1; PID:g16502823;

C.Genetics:

A.Gene: tctra

Query Match 4.4%; Score 97; DB 2; Length 1020;
Best Local Similarity 21.7%; Pred. No. 8.5; Indels 144; Gaps 19;
Matches 87; Conservative 46; Mismatches 124; Indels 144; Gaps 19;

QY 22 ISGAMPFMHGLVVAQ-----WRDYLGQ-QPDKLPIAPPTIAL-AAQFRDEIVLLG 73

DB 530 LSG--GVFVGCGKFGVSDGRYNNNSFAKVKPESGLIASKTVKXSEBRDKIA--G 585

QY 74 LKARRPVG--NHRVERISQEVAAAGLEFYGN--RRMLEKSGFPAQPPLETVAVRKVK 128

DB 586 GQSPYPAPAPWYPFVAGQLTLELTSALRGYPPLKAWISNMSNPYGVPGFLRAVAEKKLK 645

QY 129 DRRRSFYRIFPS-----GFPHGEPSQOMLSTANNREYA 166

DB 646 DRRRLPFLTADAFNMTTALADYVDPDTHNPFESGFTAPMGVASK-----A 693

QY 167 LILRHEBRPWLVCVGTENGRAPLDLAVF--RAMKLHDELGINIVPVLPMHGRPGQ 223

DB 694 TTAARPVAV--AHRRTADQ-PVSMFAFCVAKRLH-----LPGFDRAIT 738

QY 224 LPKGAVPFGEVDLDVHGTAQAVWDIRRLSWINSQEBSLIGLNGLSGGYIASVLSL 283

DB 739 DPGGNTFPLNRAEDFYLRVAANIAPWGTVPVALANQBDISLTGVSR----- 785

Qy 284 EBGACALIGVAVDLIELLGRHCGLRHKDPRRHVTXMAEPIGRMISPLSLTPLVMPGR 343
 Db 766 -----LPAI-----QHTLK-ADEVGRV-----A 802
 Qy 344 FTY-----AGIADLVHPREQVTRLMEHWKPEIWM 374
 Db 803 FTYSGRGAPRPAEDSGYTEQ-----RLGNAMWKKPLQIWM 834

RESULT 14

F71404
 Hypothetical protein - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 A:Variety: Columbia
 C:Date: 03-Aug-1998 #sequence_revision 03-Aug-1998 #ext_change 09-Jul-2004
 C:Accession: F71404
 R:Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; Dirk
 P.; Medler, H.; Medler, E.; Wambutt, R.; Weitzemegger, T.; Pohl, T.M.; Terry, N.; Giel
 vanagh, T.; Hempel, S.; Kotter, P.; Ertlan, K.D.; Rieger, M.; Schaeffer, M.; Funk, B.
 Nature 391, 485-488, 1998
 A:Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Puigdomenech
 erhof, A.; Moores, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Rechman, S.; And
 C.; Chalatzis, N.
 A:Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis thal
 A:Reference number: A71400; MUID:98121113; PMID:9461215
 A:Accession: F71404
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-505 <BEV>
 A:Cross-References: UNIPROT:Q23287, UNIPARC:UPI00000A98D0, GB:Z97335, NID:G2244747, PID:
 C:Genetics:
 A:Map position: 4COP9-4G3845

Query Match 4.4%; Score 96; DB 2; Length 505;
 Best Local Similarity 27.5%; Pred. No. 4.2;

Matches 39; Conservative 20; Mismatches 53; Indels 30; Gaps 7;

Qy 169 LNHPRPWLVCV---HGTMGAPLDLAVFRAMKLDLGLNIVPVLPMHGRGQGLP 225
 Db 105 VERPPDRP-LPCVICHGSGRADASBAIVL-----LPSNITITFLDF---SGSGLS 154
 Qy 226 KGAVPGEEDVLDDVHGTQAQW---DIRLLSWIRSQEBSLIGLNGLSLGGYIASLVA 281
 Db 155 GGE-----HYTLG--WNKXKDLKAVVEYLRTDGNVSLIGLGRSGAVTSLMYG 201
 Qy 282 SLSEGLACALIGVAVDLIELL 303
 Db 202 AEDPSIAWVLDSPFSDLVLM 223

RESULT 15

H84060
 Lysophospholipase BH3288 (imported) - Bacillus halodurans (strain C-125)

C:Species: Bacillus halodurans

C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #ext_change 12-Jul-2004

C:Accession: H84060

R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hira

Nucleic Acids Res. 28, 4317-4331, 2000

A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and

A:Reference number: AB3650; MUID:20512582; PMID:11058132

A:Accession: H84060

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-260 <STO>

A:Cross-References: UNIPROT:Q9K7S1, UNIPARC:UPI00000C415D, GB:AP01518, GB:BA000004, NID

A:Experimental source: strain C-125

C:Genetics:

A:Gene: BH3288

C:Superfamily: trophinesterase

Query Match 4.3%; Score 95; DB 2; Length 260;
 Best Local Similarity 25.5%; Pred. No. 2.1;

Matches 56; Conservative 27; Mismatches 69; Indels 68; Gaps 12;
 Qy 173 EPRPWLVCVHGT--EMGRAPLDLAVFRAM--KLHDELGLNIVPVLPMHGRGQGLPKGA 228
 Db 8 EPRGVVVVHHGAGEHNR-----YQWLAKKFNISIGLSVVMGDLF-----GOORTGX 54
 Qy 229 VFPGEEDVLDDVHGTQAQWMD--IRLLSWIRSQEBSL--IGLNGLSLGGYIASLVALEEG 286
 Db 55 -----RGHQSFOQYIDVLEWVEAKLEHVPILFLGHSMGLVA--VRTMIEG 101
 Qy 287 ---LACALIGVAVDLIELLGRHCGLRHKDPRRHVTXMAEPIGRMISPLSLTPLVMPGR 342
 Db 102 GTLPVRAVILSSPCFDLYQSPKGLASKMLHVTPTFSH----- 142
 Qy 343 RFIYAGIADRLVHPREQVTRLMEHWKPEI-----VMY 375
 Db 143 ---HSGIRSDLVTRNEIR---EAVLDELAVTKVSTKMV 176

Search completed: March 23, 2006, 05:26:36
 Job time : 17.4313 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: March 23, 2006, 04:56:43 ; Search time 105.63 Seconds
(without alignments)
2758.529 Million cell updates/sec

Title: US-10-617-038-29
Perfect score: 2195
Sequence: 1 MASSASDGTHERSAPRLSP.....AALEOSGLDAPRTQDRSA 413

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues
Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot_05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2195	100.0	413	2	006185 MYCTU
2	2188	99.7	413	2	07D6V3 MYCTU
3	2188	99.7	413	2	07Y7J1 MYCBO
4	328.5	15.0	373	2	073XK4 MYCPA
5	295.5	13.5	339	2	073TC9 MYCPA
6	293.5	13.4	400	2	08VKN4 MYCTU
7	293.5	13.4	400	2	053697 MYCTU
8	293.5	13.4	400	2	07U2C3 MYCBO
9	208.5	9.5	325	2	065J72 BACILU
10	178	8.1	333	2	081CM1 BACCR
11	158	7.2	326	2	084FB1 MYXXA
12	153.5	7.0	375	2	09C7C5 ARATH
13	148	6.7	335	2	09LH51 ARATH
14	148	6.7	360	2	09CTD8 ARATH
15	148	6.7	363	2	08RWA3 ARATH
16	146.5	6.7	366	2	06HTQ9 ORYSA
17	139.5	6.1	615	2	07NNY4 GLOVI
18	132	6.0	359	2	04NZ14 9DELTA
19	126	5.7	582	1	APBH ABRPE
20	121	5.5	338	2	06462 DROME
21	118	5.4	745	2	06F317 9PSRD
22	117.5	5.4	248	2	064UC0 BACRA
23	116.5	5.3	368	2	06C7Y6 YARLI
24	116	5.3	574	2	073S12 MYCPA
25	115.5	5.3	248	2	05LD97 BACFN
26	114	5.2	741	2	P95782 XANMA
27	114	5.2	1858	2	0939V9 9PSRU
28	112.5	5.1	1322	2	08PR21 XANAC
29	112	5.1	832	2	098G05 RHILQ
30	111	5.1	222	2	087CM7 XYLEPT
31	110.5	5.0	1925	2	05J1Q6 YNOCA

32	110	5.0	310	2	Q5WIX0 BACSK	Q5WIX0 bacillus c1
33	109	5.0	314	2	Q9XOP3 THEMA	Q9XOP3 thermotoga
34	109	5.0	502	2	Q8GWC9 ARATH	Q8GWC9 arabidopsis
35	108.5	4.9	324	2	Q6F174 MESFL	Q6F174 mesoplasma
36	108.5	4.9	751	2	Q4UPD3 XANCP	Q4UPD3 xanthomonas
37	108.5	4.9	751	2	Q8B3V6 XANCP	Q8B3V6 xanthomonas
38	108	4.9	338	2	Q84H72 BURMA	Q84H72 burkholderi
39	108	4.9	784	2	Q604E5 METCA	Q604E5 methylococ
40	107	4.9	374	2	Q721H2 THEET2	Q721H2 thetmus the
41	107	4.9	757	2	Q8PFD7 XANAC	Q8PFD7 xanthomonas
42	106.5	4.9	389	2	Q54287 STRHY	Q54287 streptomyce
43	106.5	4.9	731	1	GLGB BORPA	Q7w614 bordetella
44	106.5	4.9	749	2	Q7NEK8 GLOVI	Q7NEK8 gloeobacter
45	106.5	4.9	7756	2	Q7OLM5 BRBBE	Q7OLM5 brevibacill

ALIGNMENTS

RESULT 1	006185 MYCTU	PRELIMINARY;	PRT;	413 AA.
ID	006185 MYCTU	PRELIMINARY;	PRT;	413 AA.
AC	006185 MYCTU	PRELIMINARY;	PRT;	413 AA.
DT	01-JUL-1997 (TrEMBLrel. 04, Created)			
DT	01-AUG-1998 (TrEMBLrel. 07, Last sequence update)			
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)			
DE	Hypothetical protein.			
GN	Ordered locus names: RV2627C;			
OS	Mycobacterium tuberculosis.			
OC	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;			
OC	Corynebacterineae; Mycobacteriaceae; Mycobacterium;			
OC	Mycobacterium tuberculosis complex.			
OX	NCBI_Taxid=1773;			
RN	[1]			
RP	NUCLEOTIDE SEQUENCE.			
RC	STRAIN=H37RV;			
RC	MEDLINE=98295987; PubMed=9634230; DOI=10.1038/31159;			
RA	Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C.M.,			
RA	Harris D.E., Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III,			
RA	Tekaia F., Badcock K., Basham D., Brown D., Chillingworth T.,			
RA	Comor R., Davies R.M., Devlin K., Feltwell T., Gentles S., Hamlin N.,			
RA	Holroyd S., Hornsby T., Jagels K., Krogh A., McLean J., Moule S.,			
RA	Murphy L.D., Oliver S., Osborne K., Quail M.A., Rajandream M.A.,			
RA	Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squares R.,			
RA	Sulston J.E., Taylor K., Whitehead S., Barrett B.G.;			
RT	"Deciphering the biology of Mycobacterium tuberculosis from the			
RT	complete genome sequence."			
RL	Nature 393:537-544 (1998).			
RL	EMBL; BX842580; CAB08637.1; -; Genomic_DNA.			
DR	PIR; B70573; B70573.			
DR	TubercuList; RV2627C; -.			
DR	GO; GO:0016787; F:hydrolase activity; IEA.			
DR	InterPro; IPR000379; Ser:esterase.			
KW	Complete proteome; Hydrolyase.			
SQ	SEQUENCE 413 AA; 46252 MW; 760893D0508523B0 CRC64;			
Query Match	100.0%; Score 2195; DB 2; Length 413;			
Best Local Similarity	100.0%; Pred. No. 1e-158;			
Matches 413; Conservative	0; Mismatches 0; Indels 0; Gaps 0;			
QY	1 MASSASDGTHERSAPRLSPVLGAMGPFMTGLVVAQSWRDYLGQDPDKLPRIAPPTAL 60			
DB	1 MASSASDGTHERSAPRLSPVLGAMGPFMTGLVVAQSWRDYLGQDPDKLPRIAPPTAL 60			
QY	61 AAQAFDEIVLGLKARPPVSNHRVFERISQVAAAGLEFYGNRWLEKPSGFFAPPPLT 120			
DB	61 AAQAFDEIVLGLKARPPVSNHRVFERISQVAAAGLEFYGNRWLEKPSGFFAPPPLT 120			
QY	121 EAAVKKVDRRSFFRIFPDSGTFPHPGPSQRLSTANNREYALLRHPEPPWLYVC 180			
DB	121 EAAVKKVDRRSFFRIFPDSGTFPHPGPSQRLSTANNREYALLRHPEPPWLYVC 180			
QY	181 VAGTEWGRAFLDLAVFAWKLDHGLNIVMVLPMHGPBGGLPKGAVFGEDVLDVHV 240			

```

Db      181 VGTGTEGRAPLDLAVFRAMKLDHDELGLNIWMPVLPMHGPGQGLPGKAVFPGEEDVLDVH 240
Qy      241 GTAQAVMDIRRLLSWIRSOEBSLIGLNGLSIGYIASLVASLEBGLACALLGVPAADLI 300
Db      241 GTAQAVMDIRRLLSWIRSOEBSLIGLNGLSIGYIASLVASLEBGLACALLGVPAADLI 300
Qy      301 ELLGRHCGLRHKDPRRHVTYKMAEPIGRMISPLSLTPVLPMPGRFTYAGIADRLVHPRQV 360
Db      301 ELLGRHCGLRHKDPRRHVTYKMAEPIGRMISPLSLTPVLPMPGRFTYAGIADRLVHPRQV 360
Qy      361 TRLMEHWKPEIYVWPGHGTGFQSRPVRRFVQALLESGLLDAPRTQRDRSA 413
Db      361 TRLMEHWKPEIYVWPGHGTGFQSRPVRRFVQALLESGLLDAPRTQRDRSA 413

```

RESULT 2

```

Q7D6V3_MYCTU
ID       Q7D6V3_MYCTU PRELIMINARY; PRT; 413 AA.
AC       05-JUL-2004 (TrEMBLrel. 27, Created)
DT       05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE       05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
GN       Hypothetical protein.
OS       OrderedLocustNames=MT2702;
OC       Mycobacterium tuberculosis.
OC       Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC       Corynebacteriaceae; Mycobacteriaceae; Mycobacterium;
OC       Mycobacterium tuberculosis complex.
OX       NCBI_TaxId=1773;
RN       [1]
RP       NUCLEOTIDE SEQUENCE.
RC       STRAIN=CDC 1551 / Oshkosh;
RX       MEDLINE=22206494; PubMed=12218036;
RA       DOI=10.1126/JB.184.19.5479-5490.2002;
RA       Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA       Peterson J.D., Deboy R.T., Dodson R.J., Gwinn M.L., Hatt D.H.,
RA       Hickey E.K., Kolonay J.F., Nelson W.C., Umayam L.A., Esmolaeva M.D.,
RA       Salzberg S.L., Delcher A., Utterback T.R., Weidman J.F., Khouri H.M.,
RA       Gill J., Mikula A., Bishai W., Jacobs W.R. Jr., Venter J.C.,
RA       Fraser C.M.;
RT       "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RT       laboratory strains.";
RL       J. Bacteriol. 184:5479-5490(2002).
DR       EMBL; AE000516; AAK47018.1; -; Genomic_DNA.
DR       TIGR; MT2702; -.
DR       GO; GO:0016787; F:hydrolase activity; IEA.
KW       InterPro; IPR000379; Ser_estrs.
KW       Hydrolyase; Hypothetical protein.
SQ       SEQUENCE 413 AA; 46153 MW; 7305EBA059FABF90 CRC64;

```

Query Match 99.7%; Score 2188; DB 2; Length 413;
 Best Local Similarity 99.8%; Pred. No. 3.5e-158;
 Matches 412; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

Qy      1 MASSASDGTHERSARLSPVLSGAMGPFMTGLVYAQSWRDYLGQOPDKLPAPRTIAL 60
Db      1 MASSASDGTHERSARLSPVLSGAMGPFMTGLVYAQSWRDYLGQOPDKLPAPRTIAL 60
Qy      61 AAQAFRDEIVLLGLKARRPVSNHRVFERISQEVAAGLEFYGNRMWLEKPSGFFAOPPLT 120
Db      61 AAQAFRDEIVLLGLKARRPVSNHRVFERISQEVAAGLEFYGNRMWLEKPSGFFAOPPLT 120
Qy      121 EVAVRKVDKRRRSFYRIFPDGFTPHGEPGSGQRLSTANNREYALLRRHPEPRMVLVC 180
Db      121 EVAVRKVDKRRRSFYRIFPDGFTPHGEPGSGQRLSTANNREYALLRRHPEPRMVLVC 180
Qy      181 VGTGTEGRAPLDLAVFRAMKLDHDELGLNIWMPVLPMHGPGQGLPGKAVFPGEEDVLDVH 240
Db      181 VGTGTEGRAPLDLAVFRAMKLDHDELGLNIWMPVLPMHGPGQGLPGKAVFPGEEDVLDVH 240
Qy      241 GTAQAVMDIRRLLSWIRSOEBSLIGLNGLSIGYIASLVASLEBGLACALLGVPAADLI 300
Db      241 GTAQAVMDIRRLLSWIRSOEBSLIGLNGLSIGYIASLVASLEBGLACALLGVPAADLI 300

```

```

Db      241 GTAQAVMDIRRLLSWIRSOEBSLIGLNGLSIGYIASLVASLEBGLACALLGVPAADLI 300
Qy      301 ELLGRHCGLRHKDPRRHVTYKMAEPIGRMISPLSLTPVLPMPGRFTYAGIADRLVHPRQV 360
Db      301 ELLGRHCGLRHKDPRRHVTYKMAEPIGRMISPLSLTPVLPMPGRFTYAGIADRLVHPRQV 360
Qy      361 TRLMEHWKPEIYVWPGHGTGFQSRPVRRFVQALLESGLLDAPRTQRDRSA 413
Db      361 TRLMEHWKPEIYVWPGHGTGFQSRPVRRFVQALLESGLLDAPRTQRDRSA 413

```

RESULT 3

```

Q7TY71_MYCBO
ID       Q7TY71_MYCBO PRELIMINARY; PRT; 413 AA.
AC       07-TY71;
DT       01-OCT-2003 (TrEMBLrel. 25, Created)
DT       01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DE       01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
GN       Hypothetical protein Mb2660c.
OS       OrderedLocustNames=Mb2660c;
OC       Mycobacterium bovis.
OC       Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC       Corynebacteriaceae; Mycobacteriaceae; Mycobacterium;
OC       Mycobacterium tuberculosis complex.
OX       NCBI_TaxId=1765;
RN       [1]
RP       NUCLEOTIDE SEQUENCE.
RC       STRAIN=AF2122/97;
RX       MEDLINE=22709107; PubMed=12788972; DOI=10.1073/pnas.1130426100;
RA       Garner T., Biglmeier K., Camus J.-C., Medina N., Mansoor H.,
RA       Pryor M., Duchoy S., Gordon S., Lacroix C., Monsemp C., Simon S.,
RA       Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA       Parhill J., Barrett B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RT       "The complete genome sequence of Mycobacterium bovis.";
RL       Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
DR       EMBL; BX248343; CAD94845.1; -; Genomic_DNA.
DR       GO; GO:0016787; F:hydrolase activity; IEA.
KW       InterPro; IPR000379; Ser_estrs.
KW       Complete proteome; Hydrolyase.
SQ       SEQUENCE 413 AA; 46153 MW; 7305EBA059FABF90 CRC64;

```

Query Match 99.7%; Score 2188; DB 2; Length 413;
 Best Local Similarity 99.8%; Pred. No. 3.5e-158;
 Matches 412; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

Qy      1 MASSASDGTHERSARLSPVLSGAMGPFMTGLVYAQSWRDYLGQOPDKLPAPRTIAL 60
Db      1 MASSASDGTHERSARLSPVLSGAMGPFMTGLVYAQSWRDYLGQOPDKLPAPRTIAL 60
Qy      61 AAQAFRDEIVLLGLKARRPVSNHRVFERISQEVAAGLEFYGNRMWLEKPSGFFAOPPLT 120
Db      61 AAQAFRDEIVLLGLKARRPVSNHRVFERISQEVAAGLEFYGNRMWLEKPSGFFAOPPLT 120
Qy      121 EVAVRKVDKRRRSFYRIFPDGFTPHGEPGSGQRLSTANNREYALLRRHPEPRMVLVC 180
Db      121 EVAVRKVDKRRRSFYRIFPDGFTPHGEPGSGQRLSTANNREYALLRRHPEPRMVLVC 180
Qy      181 VGTGTEGRAPLDLAVFRAMKLDHDELGLNIWMPVLPMHGPGQGLPGKAVFPGEEDVLDVH 240
Db      181 VGTGTEGRAPLDLAVFRAMKLDHDELGLNIWMPVLPMHGPGQGLPGKAVFPGEEDVLDVH 240
Qy      241 GTAQAVMDIRRLLSWIRSOEBSLIGLNGLSIGYIASLVASLEBGLACALLGVPAADLI 300
Db      241 GTAQAVMDIRRLLSWIRSOEBSLIGLNGLSIGYIASLVASLEBGLACALLGVPAADLI 300
Qy      301 ELLGRHCGLRHKDPRRHVTYKMAEPIGRMISPLSLTPVLPMPGRFTYAGIADRLVHPRQV 360
Db      301 ELLGRHCGLRHKDPRRHVTYKMAEPIGRMISPLSLTPVLPMPGRFTYAGIADRLVHPRQV 360
Qy      361 TRLMEHWKPEIYVWPGHGTGFQSRPVRRFVQALLESGLLDAPRTQRDRSA 413
Db      361 TRLMEHWKPEIYVWPGHGTGFQSRPVRRFVQALLESGLLDAPRTQRDRSA 413

```

```

RESULT 4
073XD4 MYCPA PRELIMINARY; PRT: 373 AA.
ID 073XD4 MYCPA PRELIMINARY; AC
AC 073XD4 MYCPA PRELIMINARY; PRT: 373 AA.
DT 05-JUL-2004 (TREMblrel. 27, Created)
DT 05-JUL-2004 (TREMblrel. 27, Last sequence update)
DE Hypothetical protein.
GN OrderedlocusNames=MAP2375C;
OS Mycobacterium paratuberculosis.
OC Bacteria; Actinobacteria; Actinomycetales;
OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium avium complex (MAC).
NCBI_TaxId=1770;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K10;
RA Li L., Bannantine J., Zhang Q., Amonsin A., Alt D., Kapur V.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AEO17235; AAS04692.1; -; Genomic DNA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR InterPro; IPR000379; Ser_estrs.
KW Complete proteome; Hydrolase.
SQ
SEQUENCE 373 AA; 40489 MW; 6477DD7D520E296A CRC64;

Query Match 15.0%; Score 328.5; DB 2; Length 373;
Best Local Similarity 29.8%; Pred. No. 1.3e-16;
Matches 109; Conservative 46; Mismatches 148; Indels 63; Gaps 13;

QY 44 LGGQPDKLPAPPTTALAAQAFDEIVLLGLKARFVSNHRYVERISOEVAAGLEFYGNR 103
DB 62 LGGSPAL--RPVDCAAA--EELALGI----- 87
QY 104 RMLEKPSGFFAPPTTEVAVKVKRRSPFRIFPD--SGPTPHGEPGSRKLTSTANN 162
DB 88 -----DGAHAAAPNLPRTSI-----RRSISAGLDIERMAFEHDPLPTTLVADGIGPA 136
QY 163 REYALLLRHPE--PRPMLVCVHGTEMGRAPLDLAVFPAWKLHDELGINIWPVLPMSGPRG 221
DB 137 RAVVHLRRRRDGRPLVWVHAGGCGT--DLLSGIDRIHHLGFINIMVPOPGHGR 195
QY 222 QGLPKGAVPEGEVDLDVHGTAAQVWDIRLLSMISQEEESLIGLNGSLGYSIAY 281
DB 196 RQWP---VYPMDDPLGNVAGMRAVSEVAAYVMASSQ--ASTLVVAGISMSGPAVALVS 250
QY 282 SLEBGLACAILGVPAVDILIELGRHCGLRHKDPRHTVK-----MAEPGRMISPLST 335
DB 251 HLERRVDAVALYTPITGLNAMIARH--LQWGPARDGPRELLESVVVTLTSVIDPLAVT 308
QY 336 PLVPMRGRITVAGIDRLVHPRQVRLMEHMKKEIVVPGHTGFFQSRPVRVQA 395
DB 309 PAPPPERRITVGAHMDRMA--MRBPANADLRMA--GQLVWYDSGHVIFSRVQRTDRF 366
QY 396 LEQSGF 401
DB 367 LADAVL 372

```

```

RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K10;
RA Li L., Bannantine J., Zhang Q., Amonsin A., Alt D., Kapur V.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AEO17240; AAS06339.1; -; Genomic DNA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR InterPro; IPR000379; Ser_estrs.
KW Complete proteome; Hydrolase; Hypothetical protein.
SQ
SEQUENCE 399 AA; 43866 MW; 6B28A680BD2C697C CRC64;

Query Match 13.5%; Score 295.5; DB 2; Length 399;
Best Local Similarity 30.7%; Pred. No. 4.5e-14;
Matches 101; Conservative 43; Mismatches 144; Indels 41; Gaps 15;

QY 95 AGLEFYGNRRMLEKSGFF--AAPP-----PLTEVAVKVDRRRSFRIFFDSGF 143
DB 84 ANLEFYAEIGAAADPDSPFPPTPEPTTSRRASPLAEVARGTVD-----NLAFDSGF 137
QY 144 TP-HGEPGSGRMLSTYANNREYALLLRHPE--PRPMLVCVHGTEMGRAPLDLAVFPAWKL 201
DB 138 TAVNPAM--RAWMSGGRNNVVRAGHWRHDDGPRPTLCVTHGF--LGSSYLANGRFSLPW 194
QY 202 HDLGLNIWMPVLPMSGPRGGLP--KGAVFGEVDLDVHGTAAQVWDIRLLSMISQ 259
DB 195 YRAGYDVLMTYLPFHGKRSERFSPSGFYAGGSGFAEAMAQAVHDFRSLIDYLRGT 254
QY 260 EESLIGLNGSLGYSIAYSLSEGLACALIGVPV--ADLIE-----LLGRHGL 309
DB 255 GVER--IALTGISLGSTSLVASVDDRLAVIPNCPVTPATLFDWPAFKLVG--LGL 311
QY 310 RHKDRPRHTVKAEPGIRMIISPLSTPLVPMRGRITVAGIDRLVHPRQVRLMEHMK 369
DB 312 RSSDISR--AQLAAGLA--YHCLTYRPVLAKRMIITGLGRMA--PPEPAVKLMQHMDR 367
QY 370 PEIWWPGGHTGFFQSRPVRVQAALQ 398
DB 368 CALHWFPGSHVLIHQDLYLRMTAFLOQ 396

RESULT 6
08VKN4 MYCTU PRELIMINARY; PRT: 400 AA.
ID 08VKN4 MYCTU PRELIMINARY; AC
AC 08VKN4 MYCTU PRELIMINARY; PRT: 400 AA.
DT 01-MAR-2002 (TREMblrel. 20, Created)
DT 01-OCT-2003 (TREMblrel. 25, Last sequence update)
DE Hypothetical protein.
GN OrderedlocusNames=MT0306;
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinomycetales;
OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium tuberculosis complex.
NCBI_TaxId=1773;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CDC 1551 / Oshkosh;
RX MEDLINE=22206494; PubMed=12218036;
DOI=10.1128/JB.184.19.5479-5490.2002;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J.D., DeBoy R.T., Dodson R.J., Gwinn M.L., Haft D.H.,
RA Hickey E.K., Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D.,
RA Salzberg S.L., Delcher A., Utterback T.R., Weidman J.F., Khouri H.M.,
RA Gill J., Mikula A., Bishai W., Jacobs W.R. Jr., Venter J.C.,
RA Frazer C.M.;
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL J. Bacteriol. 184:5479-5490(2002).
DR EMBL; AB000516; AAK44530.1; -; Genomic DNA.
DR TIGR; MT0306; -;
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR InterPro; IPR000379; Ser_estrs.
KW Hydrolase; Hypothetical protein.

```

SEQ	SEQUENCE	400 AA;	44011 MW;	D6FD9510B3ABB6CB8 CRC64;
QY	Query Match	13.4%;	Score 293.5;	DB 2; Length 400;
DB	Best Local Similarity	30.5%;	Pred. No. 6.5e-14;	
Matches	97; Conservative	42;	Mismatches 122;	Indels 57; Gaps 14
QY	95 AGLEFYGNRRWLEKSPGFPAPQ-----PULTEVAARKVQDRRRSFYRIFPDGCF	143		
DB	84 SNLNFYAEILAAAHDAKSPFAPTELPKVTSRPASPLTEWVARGTVD-----NIAFASGF	137		
QY	144 TP-HGGEPPSCQWLVSTANNRRYALLNRP-PRPLVCVGHTEWGRAPLDLAVRAMKL	201		
DB	138 RAINP-TWRQKWSALTANNIVHAQWHRHRDPRPLCVIHGF-MGSSYLNLGLFPSPWP	194		
QY	202 HDELGLNIVPVLPMHGPFGQGLPKGAVPPG-----EDVLDVHGTGAQVMDIRLLSWI	256		
DB	195 YRSRGVDVLLYTLPPHGGQAE---KFSPPSGGCFYFSGLSGRFAEMAAQVYDFRSIVDTL	251		
QY	257 RSQEEBSLGLNGLSLGTYIASLVASLSEGLACAILGVPA-----DLIELL	303		
DB	252 RHIGVDR-IALTGISLGGYTSALTLSVESRLEAVIPNCPVWPAKLFDPEWPKXKLVLK-	309		
QY	304 GNHGLRRKDRPRRHVYKMAEPIGRMI--SPSLTLPLVMPHGRFYAGIADRLVHREQVT	361		
DB	310 ---GLRLTN-----ISRDELLAGLAVYMGPLVYRLLPRDRMTITGLDRMAPEHAAVT	360		
QY	362 RLMEHWKREIYVWPGGH	379		
DB	361 -LWKQMDRCALHWPFGSH	377		
RESULT 7	053697 MYCTU PRELIMINARY; PRT; 400 AA.			
ID	053697 MYCTU PRELIMINARY; PRT; 400 AA.			
AC	053697			
DT	01-JUN-1998 (TREMBLrel. 06, Created)			
DT	01-JUN-1998 (TREMBLrel. 06, Last sequence update)			
DT	01-OCT-2003 (TREMBLrel. 25, Last annotation update)			
DE	Hypothetical Protein.			
GN	OrderedCusNames=Rv0293c;			
OS	Mycobacterium tuberculosis.			
OC	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;			
OC	Corynebacteriineae; Mycobacteriaceae; Mycobacterium;			
OC	Mycobacterium tuberculosis complex.			
OX	NCBI_TaxId=1773;			
RN	[1]			
RP	NUCLEOTIDE SEQUENCE.			
RC	STRAIN=H37RV;			
RC	MEDLINE=98295987; PubMed=9634230; DOI=10.1038/11159;			
RA	Cole S.T., Brosch R., Parkhill J., Garnier T., Church C.M.,			
RA	Harris D.E., Gordon S.V., Bignellier K., Gas S., Barry C.E. III,			
RA	Tekaia F., Badcock K., Basham D., Brown D., Chillingworth T.,			
RA	Connor R., Davies R.M., Devlin K., Feltham T., Gentles S., Hamlin N.,			
RA	Holroyd S., Hornbly T., Jagels K., Krogh A., McLean J., Moutle S.,			
RA	Murphy L.D., Oliver S., Osborne J., Quail M.A., Ratandream M.A.,			
RA	Rogers J., Rutter S., Seeger K., Skellon S., Squares S., Squares R.,			
RA	Sulston J.E., Taylor K., Whitehead S., Barrett B.G.;			
RT	"Deciphering the biology of Mycobacterium tuberculosis from the			
RT	complete genome sequence."			
RL	Nature 393:537-544 (1998).			
RL	EMBL; BX842573; CAA17368.1; -; Genomic_DNA.			
DR	PIR; C70837; C70837.			
DR	TubercuList; Rv0293c; -			
DR	GO; GO:0016787; F:hydrolyase activity; IEA.			
DR	InterPro; IPR000379; Ser_Peptide.			
KW	Complete proteome; Hydrolyase; Hypothetical protein.			
SQ	SEQUENCE 400 AA; 43971 MW; FBBBC136ADFAABC CRC64;			
QY	Query Match	13.4%;	Score 293.5;	DB 2; Length 400;
DB	Best Local Similarity	30.5%;	Pred. No. 6.5e-14;	
Matches	97; Conservative	42;	Mismatches 122;	Indels 57; Gaps 14
QY	95 AGLEFYGNRRWLEKSPGFPAPQ-----PULTEVAARKVQDRRRSFYRIFPDGCF	143		
DB	84 SNLNFYAEILAAAHDAKSPFAPTELPKVTSRPASPLTEWVARGTVD-----NIAFASGF	137		
QY	144 TP-HGGEPPSCQWLVSTANNRRYALLNRP-PRPLVCVGHTEWGRAPLDLAVRAMKL	201		
DB	138 RAINP-TWRQKWSALTANNIVHAQWHRHRDPRPLCVIHGF-MGSSYLNLGLFPSPWP	194		
QY	202 HDELGLNIVPVLPMHGPFGQGLPKGAVPPG-----EDVLDVHGTGAQVMDIRLLSWI	256		
DB	195 YRSRGVDVLLYTLPPHGGQAE---KFSPPSGGCFYFSGLSGRFAEMAAQVYDFRSIVDTL	251		
QY	257 RSQEEBSLGLNGLSLGTYIASLVASLSEGLACAILGVPA-----DLIELL	303		
DB	252 RHIGVDR-IALTGISLGGYTSALTLSVESRLEAVIPNCPVWPAKLFDPEWPKXKLVLK-	309		
QY	304 GNHGLRRKDRPRRHVYKMAEPIGRMI--SPSLTLPLVMPHGRFYAGIADRLVHREQVT	361		
DB	310 ---GLRLTN-----ISRDELLAGLAVYMGPLVYRLLPRDRMTITGLDRMAPEHAAVT	360		
QY	362 RLMEHWKREIYVWPGGH	379		
DB	361 -LWKQMDRCALHWPFGSH	377		

DB	Seq	Score	Length	DB 2	Length	DB 3	Length	DB 4	Length	DB 5	Length	DB 6	Length	DB 7	Length	DB 8	Length	DB 9	Length	DB 10	Length	DB 11	Length	DB 12	Length	DB 13	Length	DB 14	Length	DB 15	Length	DB 16	Length	DB 17	Length	DB 18	Length	DB 19	Length	DB 20	Length	DB 21	Length	DB 22	Length	DB 23	Length	DB 24	Length	DB 25	Length	DB 26	Length	DB 27	Length	DB 28	Length	DB 29	Length	DB 30	Length	DB 31	Length	DB 32	Length	DB 33	Length	DB 34	Length	DB 35	Length	DB 36	Length	DB 37	Length	DB 38	Length	DB 39	Length	DB 40	Length	DB 41	Length	DB 42	Length	DB 43	Length	DB 44	Length	DB 45	Length	DB 46	Length	DB 47	Length	DB 48	Length	DB 49	Length	DB 50	Length	DB 51	Length	DB 52	Length	DB 53	Length	DB 54	Length	DB 55	Length	DB 56	Length	DB 57	Length	DB 58	Length	DB 59	Length	DB 60	Length	DB 61	Length	DB 62	Length	DB 63	Length	DB 64	Length	DB 65	Length	DB 66	Length	DB 67	Length	DB 68	Length	DB 69	Length	DB 70	Length	DB 71	Length	DB 72	Length	DB 73	Length	DB 74	Length	DB 75	Length	DB 76	Length	DB 77	Length	DB 78	Length	DB 79	Length	DB 80	Length	DB 81	Length	DB 82	Length	DB 83	Length	DB 84	Length	DB 85	Length	DB 86	Length	DB 87	Length	DB 88	Length	DB 89	Length	DB 90	Length	DB 91	Length	DB 92	Length	DB 93	Length	DB 94	Length	DB 95	Length	DB 96	Length	DB 97	Length	DB 98	Length	DB 99	Length	DB 100	Length	DB 101	Length	DB 102	Length	DB 103	Length	DB 104	Length	DB 105	Length	DB 106	Length	DB 107	Length	DB 108	Length	DB 109	Length	DB 110	Length	DB 111	Length	DB 112	Length	DB 113	Length	DB 114	Length	DB 115	Length	DB 116	Length	DB 117	Length	DB 118	Length	DB 119	Length	DB 120	Length	DB 121	Length	DB 122	Length	DB 123	Length	DB 124	Length	DB 125	Length	DB 126	Length	DB 127	Length	DB 128	Length	DB 129	Length	DB 130	Length	DB 131	Length	DB 132	Length	DB 133	Length	DB 134	Length	DB 135	Length	DB 136	Length	DB 137	Length	DB 138	Length	DB 139	Length	DB 140	Length	DB 141	Length	DB 142	Length	DB 143	Length	DB 144	Length	DB 145	Length	DB 146	Length	DB 147	Length	DB 148	Length	DB 149	Length	DB 150	Length	DB 151	Length	DB 152	Length	DB 153	Length	DB 154	Length	DB 155	Length	DB 156	Length	DB 157	Length	DB 158	Length	DB 159	Length	DB 160	Length	DB 161	Length	DB 162	Length	DB 163	Length	DB 164	Length	DB 165	Length	DB 166	Length	DB 167	Length	DB 168	Length	DB 169	Length	DB 170	Length	DB 171	Length	DB 172	Length	DB 173	Length	DB 174	Length	DB 175	Length	DB 176	Length	DB 177	Length	DB 178	Length	DB 179	Length	DB 180	Length	DB 181	Length	DB 182	Length	DB 183	Length	DB 184	Length	DB 185	Length	DB 186	Length	DB 187	Length	DB 188	Length	DB 189	Length	DB 190	Length	DB 191	Length	DB 192	Length	DB 193	Length	DB 194	Length	DB 195	Length	DB 196	Length	DB 197	Length	DB 198	Length	DB 199	Length	DB 200	Length	DB 201	Length	DB 202	Length	DB 203	Length	DB 204	Length	DB 205	Length	DB 206	Length	DB 207	Length	DB 208	Length	DB 209	Length	DB 210	Length	DB 211	Length	DB 212	Length	DB 213	Length	DB 214	Length	DB 215	Length	DB 216	Length	DB 217	Length	DB 218	Length	DB 219	Length	DB 220	Length	DB 221	Length	DB 222	Length	DB 223	Length	DB 224	Length	DB 225	Length	DB 226	Length	DB 227	Length	DB 228	Length	DB 229	Length	DB 230	Length	DB 231	Length	DB 232	Length	DB 233	Length	DB 234	Length	DB 235	Length	DB 236	Length	DB 237	Length	DB 238	Length	DB 239	Length	DB 240	Length	DB 241	Length	DB 242	Length	DB 243	Length	DB 244	Length	DB 245	Length	DB 246	Length	DB 247	Length	DB 248	Length	DB 249	Length	DB 250	Length	DB 251	Length	DB 252	Length	DB 253	Length	DB 254	Length	DB 255	Length	DB 256	Length	DB 257	Length	DB 258	Length	DB 259	Length	DB 260	Length	DB 261	Length	DB 26
----	-----	-------	--------	------	--------	------	--------	------	--------	------	--------	------	--------	------	--------	------	--------	------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	-------

OK NCBI_TaxID=34;
 (1)
 NUCLEOTIDE SEQUENCE.
 RA Hartzell P.L., Youderian P.A.;
 RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY04469; M022902.1; -; Genomic_DNA.
 KW Hypothetical protein.
 SQ SEQUENCE 326 AA; 35332 MW; C82D125BA41518DB CRC64;

Query Match 7.2%; Score 158; DB 2; Length 326;
 Best Local Similarity 24.6%; Pred. No. 0.0011;
 Matches 88; Conservative 36; Mismatches 138; Indels 96; Gaps 17;

QY 68 EIVLLGLKARRPVSNNRVF-----ERISQEVAAGLEFYGN-----RRMLEK 108
 DB 8 DVLPLGLSRR-----RLFSQGWGDEQPLEEVAAAPFGQRRPPLAPEMSAVRLQGLQV 62
 QY 109 PGGFPAQPPPLTEVAVRKXKRRRSFYRLFPDSGFTPHGEGSQWLS----- 157
 DB 63 RDGTFPSPLARLDAAAR-----TAH-----VMTLSAGQSPSGACV 98
 QY 158 YANNREYVALLRHPRPFWLVCVSTGEMGRAPLDLAVRAWKHDELGLNTVMPVLPWH 217
 DB 99 VLAASREESGSLRERYAP-----IARSGIDLFL-----LENPYGGR 136
 QY 218 GPRGQGLPKGAVFPGEDVDVHGTAAVMDIRRLLSWRSQEEESLIGLNGLSIGYIA 277
 DB 137 RVVGQ---KGAL--RTVSDHVLMLGWDVARALLAMRS--EGHARLGVAGYSGWGYMA 190
 QY 278 SLIVAS-LEEGALCALIGV--PYADLIE-LGHHGCLRKND-PRRTTVKMAEPICGMISIP 331
 DB 191 ALTAAVVPEPLVVAALAGASPVVPTQGLSWSIAFALLDGRRAQDAQSRRLGIFDL 250
 QY 332 LSLTLPVP--MPGRFYAGIADRLVPRBOVTRLMEHWKKEPIVWYPGGH-TGFFOSR 386
 DB 251 ANLAPPPRQPEAAVLVACRRDGVFGDETTALAHNMPRSELRWVDAQHTALPTFR 308

RESULT 12
 Q9C7C5 ARATH PRELIMINARY; PRT; 375 AA.

AC Q9C7C5;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein T21B14.3.
 GN GN Name: T21B14.3;
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 OC eurosida II; Brassicales; Brassicaceae; Arabidopsis.
 OK NCBI_TaxID=3702;

RP NUCLEOTIDE SEQUENCE.
 RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.J., Wu D.,
 RA Maiti R., Roming C.M., Koo H., Fujii C.Y., Uteback T.R.,
 RA Barstead M.E., Bowman C.L., White O., Niernan W.C., Fraser C.M.;
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC069473; ANG51056.1; -; Genomic_DNA.
 KW Hypothetical protein.
 SQ SEQUENCE 375 AA; 42136 MW; A36E7F425875B6A2 CRC64;

Query Match 7.0%; Score 153.5; DB 2; Length 375;
 Best Local Similarity 28.0%; Pred. No. 0.0028;
 Matches 97; Conservative 40; Mismatches 131; Indels 79; Gaps 21;

QY 82 NHRVERISQEVAAAGLEFYGNRRWLEKPSGFFAQPPLTEVAVRKVKDRRSFYR--IFF 139
 DB 39 NIELLERVQRLFP-LEVOG-QNW-----PPLVPRVWRTVMEKTKATLREGVF- 85
 QY 140 DSGFTPHGE-----PGSQRWLSYANNREYVALLRHPRPFWLVCVSTGEMGRAPL-DL 193
 DB 86 ---QTPCADELTAALPPEBSR---TA--RVAMLVPRKVPQKMACVHLAVVARACLCDL 136

QY 194 AVF-----RAWKHDEL-GLNTVMPVL--PMHPRGQGLPKGA-VPRGEDVLDV 239
 DB 137 NLFIPLPDHYYDRRLRGGLVKNINATWVLESFFYQRRPPLQCGRRLLCVSDLL--L 194
 QY 240 HGTAAVMDIRRLLSWRSQEEESLIGLNGLSIGYIASLVASLEGALCALIGVAVDL 299
 DB 195 LGRA-TIESRSLIHWLDTBEGFGKMGVCGLSMGVHAASVGSLSHP-----TPVATL 245
 QY 300 IELLGRH-----C-----GLRHK-DPRRHTVKAEPICGMISPLSLT-----P 336
 DB 246 -PFLSPHAAVAFCEGLIKGTAWMALREELAAQKITWLTDEVERKMNVLSTDTVRFP 304
 QY 337 LVPMGFRFYAGIADRLVPRBOVTRLMEHWKKEPIVWYPGGHGTF 383
 DB 305 IPKNDAVIFVAATDDGYIPKHSVLELQKAMPGESEVRWVTGGHVSF 351

RESULT 13
 Q9LH51 ARATH PRELIMINARY; PRT; 335 AA.

AC Q9LH51;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Similarity to unknown protein.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 OC rosids; eurosida II; Brassicales; Brassicaceae; Arabidopsis.
 OK NCBI_TaxID=3702;

RN (1)
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=2036309; PubMed=10907853;
 RA Kaneko T., Kato T., Sato S., Nakamura Y., Asamizu E., Tabata S.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 3. II.
 RT Sequence features of the 4,251,695 bp regions covered by 90 pl, TAC
 RT and BAC clones.";
 RL DNA Res. 7:217-221 (2000).
 RN (2)

RP NUCLEOTIDE SEQUENCE.
 RA Kaneko T., Kato T., Sato S., Nakamura Y., Asamizu E., Tabata S.;
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AP002063; BAB01965.1; -; Genomic_DNA.
 SQ SEQUENCE 335 AA; 37450 MW; D8AD777D6F11C27A CRC64;

Query Match 6.7%; Score 148; DB 2; Length 335;
 Best Local Similarity 27.7%; Pred. No. 0.0064;
 Matches 93; Conservative 38; Mismatches 133; Indels 72; Gaps 20;

QY 82 NHRVERISQEVAAAGLEFYGNRRWLEKPSGFFAQPPLTEVAVRKVKDRRSFYR--IFF 139
 DB 14 NIELLERVQRLFP-LEVOG-QNW-----PPLVPRVWRTVMEKTKATLREGVF- 60
 QY 140 DSGFTPHGE-----PGSQRWLSYANNREYVALLRHPRPFWLVCVSTGEMGRAPLDA 194
 DB 61 ---QTPCADELTAALPPEBSR---TA--RVAMLVPRKVPQKMACVHAAGTGHYD-- 109
 QY 195 VFRANKHDEL-GLNTVMPVL--PMHPRGQGLPKGA-VPRGEDVLDVHGTAAVMDIR 250
 DB 110 --RRLRGGLVKNINATWVLESFFYQRRPPLQCGARRLLCVSDLL--LGRA-TIESR 164
 QY 251 RLISWRSQEEESLIGLNGLSIGYIASLVASLEGALCALIGVAVDLIELGHN----- 306
 DB 165 SLIHLDTBEGRGKRGVCGLSMGVHAASVGSLSHP-----TEVATL-PFLSPHAAV 215
 QY 307 --C-----GLRHK-DPRRHTVKAEPICGMISPLSLT-----PLVMPGRFYA 347
 DB 216 AFCEGILKGTAWMALREELAAQKITWLTDEVERKMNVLSTDTVRPPIKNDPAVIFV 275
 QY 348 GIADRLVPRBOVTRLMEHWKKEPIVWYPGGHGTF 383
 DB 276 AATDGYIPKHSVLELQKAMPGESEVRWVTGGHVSF 311

RESULT 14

O9C7D8 ARATH

ID O9C7D8 ARATH PRELIMINARY; PRT: 360 AA.

AC O9C7D8; 01-JUN-2001 (TrEMBLrel. 17, Created)

DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Hypothetical protein F28J15.20.

GN Name=F28J15.20;

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;

OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.

OX NCBI_Taxid=3702;

RN [1] NUCLEOTIDE SEQUENCE.

RP Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.J., Wu D.,

RA Malet R., Rongming C.M., Koo H., Fujii C.Y., Utecherback T.R.,

RA Barnstead M.E., Bowman C.L., White O., Nierman W.C., Fraser C.M.;

RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AC069472; AAC51070.1; -; Genomic_DNA.

KW Hypothetical protein.

SQ SEQUENCE 360 AA; 40396 MW; EA53832CD3C404C CRC64;

Query Match 6.7%; Score 148; DB 2; Length 360;

Best Local Similarity 27.7%; Pred. No. 0.0069;

Matches 93; Conservative 38; Mismatches 133; Indels 72; Gaps 20;

82 NHRVERISQEVAAAGLEFYGNRRMLEKPSGFFAQPPLTEVAVRKVKDRRSFYR--IFP 139

39 NLELERNVQRLFP--LEVQG-QNW-----PPPLVRPVWRTWETKTATLREGVF- 85

140 DSGFTPHPE-----PGSORMLSYTANNREYALLRHPPRPMLVCVHGTEMGRAPLDLA 194

86 ---QTPCADELTAALPPESR---TA--RVAMLVKKNVPPQKACVHLAAGTGDHTYD-- 134

195 VERAWKLDEL--GLNIIVPVL--PMHGPRGQGLPKGA--VFGEDVLDVHGTAQAVMDIR 250

135 --RLRLGGPLVKQNIATVLESFPGQRRPFLQCGARLLCVSDLL--LLGRA-TIEESR 189

251 RLWSIRSQEESLLGLNGSLGGYIASLVASLEGLACAILGVPAADIIEILGRH---- 306

190 SLIHWLDTBEFGKMGVCGLSMGVHVASVGLHP-----TPVATL--PFLSPHSAVV 240

307 --C-----GLRHK-DPRRHTVKAEPIGRMISPLSLT-----PLVMPGRTFYA 347

241 AECGELIKKGTAWMALREBLAAQKITMTLDEVEREMRNVLSTDVTRPPIPKNPDAVIFV 300

348 GIADRLVHPREQVTRLMWHMGKPEIWMYPGGHTGFF 383

301 AATDGYIPKHSVLELQKAMPSEVRAWVTGCHVSSF 336

Db

Qy

Qy

Qy

Qy

Qy

Qy

Qy

Qy

Qy

Qy

Qy

Qy

Qy

Qy

Qy

Qy

Qy

Qy

Qy

Qy

Qy

Qy

Qy

Qy

Qy

Qy

Qy

Qy

RA Cheuk R., Chung M.K., Hayaehizaki Y., Ishida J., Kamiya A., Kawai J.,

RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,

RA Sakurai T., Satou M., Seki M., Shim P., Yamada K., Shinozaki K.,

RA Ecker J., Theologis A., Davis R.W.;

RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.

RN [2]

RP NUCLEOTIDE SEQUENCE.

RA Kim C.J., Chen H., Cheuk R., Shim P., Bowser L., Carninci P.,

RA Dale J.M., Hayaehizaki Y., Ishida J., Jones T., Kamiya A.,

RA Karlin-Neumann G., Kawai J., Lam B., Lin J., Miranda M., Narusaka M.,

RA Nguyen M., Onodera C.S., Palm C.J., Quach H.L., Sakurai T., Satou M.,

RA Seki M., Southwick A., Toriumi M., Wong C., Wu H.C., Yamada K., Yu G.,

RA Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;

RA Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.

RL EMBL; AY093234; AAM13233.1; -; mRNA.

DR EMBL; BT008851; AAP68290.1; -; mRNA.

KW Hypothetical protein.

SQ SEQUENCE 363 AA; 40784 MW; 6A86B799868E1748 CRC64;

Query Match 6.7%; Score 148; DB 2; Length 363;

Best Local Similarity 27.7%; Pred. No. 0.007;

Matches 93; Conservative 38; Mismatches 133; Indels 72; Gaps 20;

82 NHRVERISQEVAAAGLEFYGNRRMLEKPSGFFAQPPLTEVAVRKVKDRRSFYR--IFP 139

42 NLELERNVQRLFP--LEVQG-QNW-----PPPLVRPVWRTWETKTATLREGVF- 88

140 DSGFTPHPE-----PGSORMLSYTANNREYALLRHPPRPMLVCVHGTEMGRAPLDLA 194

89 ---QTPCADELTAALPPESR---TA--RVAMLVKKNVPPQKACVHLAAGTGDHTYD-- 137

195 VERAWKLDEL--GLNIIVPVL--PMHGPRGQGLPKGA--VFGEDVLDVHGTAQAVMDIR 250

138 --RLRLGGPLVKQNIATVLESFPGQRRPFLQCGARLLCVSDLL--LLGRA-TIEESR 192

251 RLWSIRSQEESLLGLNGSLGGYIASLVASLEGLACAILGVPAADIIEILGRH---- 306

193 SLIHWLDTBEFGKMGVCGLSMGVHVASVGLHP-----TPVATL--PFLSPHSAVV 243

307 --C-----GLRHK-DPRRHTVKAEPIGRMISPLSLT-----PLVMPGRTFYA 347

244 AECGELIKKGTAWMALREBLAAQKITMTLDEVEREMRNVLSTDVTRPPIPKNPDAVIFV 303

348 GIADRLVHPREQVTRLMWHMGKPEIWMYPGGHTGFF 383

304 AATDGYIPKHSVLELQKAMPSEVRAWVTGCHVSSF 339

Db

Qy

Qy

Qy

Qy

Qy

Qy

Qy

Qy

Qy

Qy

Qy

Qy

Qy

Qy

Qy

Qy

Qy

Qy

Qy

Qy

Qy

Qy

Qy

Qy

Qy

Qy

Qy

Qy

Qy

Qy

Qy

Qy

Qy

This Page Blank (uspto)

;; PRIOR APPLICATION NUMBER: US 60/117,747
;; PRIOR FILING DATE: 1999-01-29
;; NUMBER OF SEQ ID NOS: 14342
;; SEQ ID NO 7241
;; LENGTH: 889
;; TYPE: PRT
;; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-7241

Query Match 5.4%; Score 117.5; DB 2; Length 889;
Best Local Similarity 22.8%; Pred. No. 0.0034;
Matches 100; Conservative 59; Mismatches 158; Indels 121; Gaps 26;

QY 20 PVLGAMG-----PFGHTGLVVAQSWRD-YIGQOPDKLPAPRTIALAQAQFRDEIVLL 72
DB 240 PVITALLNNAAGFIKIVIRGQL-----SPADVFPDDPSKKP-----RBAQSFIDQVGP 289
QY 73 GL-----KARRVSNHRVPERISQEVAAQLFEYGNRRMLKESGFFAOPPLTEVAV 124
DB 290 GLYLLEAQMGMKTEBAALYAAYOMLVQEKATGI-----YFALPTQULTS--- 332
QY 125 RKVKORRRSFYRIFPDGFTPHGEPGSGQWMLSTYANNREYALLRHDEPRMPLVCVHGT 184
DB 333 NKIDRFNAFLHQIV-STETP-----QHSLL--HSGAKMLM---DT 367
QY 185 EMGRAPLDLAVFRWKLDHDELGL-----NIVMPVLPW-HG-PRGQGLPKGAVFPG 232
DB 368 EMGE---EGSPGGAWFNRKRKGLLAPFVGTIDQALMVMVNVKHGFVAYGLAKGV--- 421
QY 233 EDVLDVH-----GT-AQAVVDIR-----LISWRSQ-EESLSLGLNLSIGYIA 277
DB 422 --LIDEVHTYDLTYGTIINALVEFLRQIDCTVILSATLSQTRDALLQOSTTSAYPLI 479
QY 278 SLVASLEEGLCALIGVPVADLIELGHRHCGLRHDPRRHTVMAEPIGRMISPLSLT-- 335
DB 480 TAASBAEREGVIEGVVTEHTYIILSC--RKQEPAREELRAELGQGVLENTIIA 537
QY 336 ---PLVMPGRFIYAGIADRLVHPR---EQVTRLMHWKPEIIVYFGHTGFQSPRV 388
DB 538 EAQGYLDLARSRAVAGCETGLHRSFTPHNRHREQW---VALY--GRAGWPFQRKC 591
QY 389 RRFV--QAALFQSGILDA 404
DB 592 GRIVGTQVLEQSLDIDA 609

RESULT 3

US-09-252-991A-20236
; Sequence 20236, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20236
; LENGTH: 409
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20236

Query Match 5.1%; Score 112.5; DB 2; Length 409;
Best Local Similarity 24.7%; Pred. No. 0.0035;
Matches 94; Conservative 33; Mismatches 111; Indels 143; Gaps 23;

QY 105 WLEKPSGFFAOPPLTEVAVRKVKORRRSFYRIFPDG---FTPHGEPGSGQWMLSY-- 158

DB 66 WDEYATGVIRRHPRAT-----ADNRR--RLPALGPPRATLEGHPG-GGSVPMEGYAT 114
QY 159 -----TANNREYALLLRHPRP-----RPMLVCV----- 181
DB 115 GVIRRYRAATADNKRKLEALQAHVPCRKATGMPFCMAATNATGAIKCLPAVLHRHPAQV 174
QY 182 ---HGTMGRAPLDLAVFRWKLDHDELGLNVN--PVL-----PMGPR-----GQ 222
DB 175 DTDHRRVVVG--GPVRLGVQAVLLPRGLELDVVVQHGLGVGVDPRLRAAYLARQAGGE 233
QY 223 GLPKGAVFPGBDV-----LDDVH-----GTQAVWDIR--RLLSWR 257
DB 234 GVVAVVAGLPGSHVIALSCARLTQHQVAGAADVHRVELGIGIEVADDQATRLIGARR 293
QY 258 SQEESLSLGLNLSIGYIASLVASLEGLACALIGVPVADLIELGHRGCLR--HKDPR 315
DB 294 IARQPVHGLGAGARPG-----QVAVVAVAGRVV---VABRALRLQVVRHDCQ 339
QY 316 RH-TVKRAEPIGRMISPLSLTPVLVMPGRFTYAGIADRLVHPRQVTRLMWH---WGP 370
DB 340 AFAAIDLLEGLGQ-----GRTV-AGVVELRVH-----RTVEHRRGSIGAD 378
QY 371 EIVWYPGHTGFF--QSRPVR 389
DB 379 LVAALDQPHGNFVAALQRLPVR 399

RESULT 4

US-09-252-991A-26849
; Sequence 26849, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 26849
; LENGTH: 1127
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26849

Query Match 5.0%; Score 109.5; DB 2; Length 1127;
Best Local Similarity 23.5%; Pred. No. 0.036;
Matches 84; Conservative 36; Mismatches 100; Indels 137; Gaps 18;

QY 123 AVRKVKORRRSFYRIFPDGFTPHGEPGSGQWMLSTYANNREYALLLRHPRP--WLV 179
DB 27 AARSIPSR-----PGRPCNR-----CWTTRSVPRLPPGSAFASAWW-- 65
QY 180 CVHGTMGRAPLDLAVFRWKLDH--DEGLNIVMPVLRPMHGRGGLPKGAV-----FPG 232
DB 66 ---SRSGRCPRDSCFTSLRFHYDOERLELRIGLIVAHGEGGIGIGQEAVALRAGARV 121
QY 233 EDVLDVHGTQAQAVWDIRRLRLSWIRSQEESL---IGNLGLSIGYIASLVASLEGLA 288
DB 122 DRDADVHGLA-----VHLQRADLGNHRRRLDVAAVGGDL-DLVA-----GID 164
QY 289 CAILGVPAVDLIELGHRHCGLRHDPRRHTVMAEPYGRMISPLSLTPV---PMGR 343
DB 165 AQLLGORLADLDELGL-----LGDRLQPAVLGVVLEVLEGPVGR 204
QY 344 FI-----YAGIADRL-----VHPREOVTRL-----WEHMGKPEI 372
DB 205 RVGLGVRAEQFHVVLVHPRKRVADRLAVVAVGVHFDRLRLRLVVLGGRAFGLVDDGQ 264

Qy 373 WYPCG-----HTGFQSRPVRRFQAL-----ECSGLDAPRTORD 410
 Db 265 ASHPGVADDERVHPGF-----RRTVGLVGVHGAAPGLAVPRAQLLRVRLAAD 315

RESULT 5

US-09-976-674-43
 ; Sequence 43, Application US/09976674
 ; Patent No. 6644180
 ; GENERAL INFORMATION:
 ; APPLICANT: Qi, Steve
 ; APPLICANT: Akinsanya, Karen
 ; APPLICANT: Riviere, Pierre
 ; APPLICANT: Junien, Jean-Louis
 ; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
 ; FILE REFERENCE: 70669
 ; CURRENT APPLICATION NUMBER: US/09/976,674
 ; CURRENT FILING DATE: 2001-10-12
 ; PRIOR APPLICATION NUMBER: US 60/240,117
 ; PRIOR FILING DATE: 2000-10-12
 ; NUMBER OF SEQ ID NOS: 61
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 43
 ; LENGTH: 691
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-976-674-43

Query Match 4.9%; Score 106.5; DB 2; Length 691;
 Best Local Similarity 18.2%; Pred. No. 0.035;
 Matches 63; Conservative 63; Mismatches 106; Indels 115; Gaps 14;

Qy 9 THNSAFRLSPVYLSGANGPFFMTGLVYAQSRDYLGQOPDKLPYARPTIALAAQAFRDE 68
 Db 288 SRDGSKFMTYVVKQSGGSEFHHIMFLIQSKSEOI-----TVRHLTSGNMEV 335
 Qy 69 IYVLGLKRRPVSNNRVRFERISOEVAAGLEFYGNRMLKEKSPGFAQPPRLREVAVRKYK 128
 Db 336 IKIL-----AYDETQKISASTEGILNRQC-----ISCNFMK 367
 Qy 129 DRRRSFYRIFFDSGFTPHGEPBSQRMWLSY-----TANNREYALL----- 168
 Db 368 EQ-----CTYFPAASPM-----NQHFLFCGSPRPVVSLSHTDNPAKYFLLENSMLK 417
 Qy 169 ---LHPBPRLVLCVHGTEMGRAPLDLAVFRAMKLHDELGLNIWPIVPMHGPRGQLP 225
 Db 418 EAILKKIKGPKRIKILHIDY-ELPLQLSLPKDFMDRQYALLIMD-----EPPGQULVT 472
 Qy 226 -----KGAVPRGEDVLDVYH---GTAQAVMDIRL-----L 253
 Db 473 DKFHIDWSVLLIDMNVIVARPDGSGSGFQGLKTIQETHRLRGSVEVKKQITAVXVFLKL 532
 Qy 254 SWIRSOEBSLIGLNGLSIGYIASLVAISLEGL-ACAILGVAVDYL 299
 Db 533 PYIDSKR-----LSIFGKGYGCIYIASMLKSDKRLPKCGSVAVPIIDL 575

RESULT 6

US-08-764-100-27
 ; Sequence 27, Application US/08764100
 ; Patent No. 5773700
 ; GENERAL INFORMATION:
 ; APPLICANT: van Ginsven J., Martine Q.
 ; APPLICANT: De Haan, Petrus T.
 ; APPLICANT: Gielen L., Johannes J.
 ; APPLICANT: Peters, Dirk
 ; APPLICANT: Goldbach, Robert W.
 ; TITLE OF INVENTION: Improvements in or Relating to Organic
 ; NUMBER OF SEQUENCES: 27
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Sandoz Agro, Inc

STREET: 975 California Avenue
 CITY: Palo Alto
 STATE: CA
 COUNTRY: USA
 ZIP: 94304

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/764,100
 FILING DATE: 06-DEC-1996
 CLASSIFICATION: 800
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/08/214,064

FILING DATE:
 APPLICATION NUMBER: US 08/032,235
 FILING DATE: 17-MAR-1993
 APPLICATION NUMBER: GB 9206016.9
 FILING DATE: 19-MAR-1992

ATTORNEY/AGENT INFORMATION:
 NAME: No. 577370019, Allen E.
 REGISTRATION NUMBER: 34,490
 REFERENCE/DOCKET NUMBER: 137-1061
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 354-3592
 TELEFAX: (415) 857-1125

INFORMATION FOR SEQ ID NO: 27:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 3218 amino acids
 TYPE: amino acid
 STRANDEDNESS: unknown
 TOPOLOGY: unknown

MOLECULE TYPE: Protein
 ; US-08-764-100-27

Query Match 4.9%; Score 106.5; DB 1; Length 3218;
 Best Local Similarity 22.4%; Pred. No. 0.39;
 Matches 81; Conservative 40; Mismatches 158; Indels 83; Gaps 14;

Qy 33 GLVYAQSRDYLGQOPDKLPYAR--PTIALAAQAFRDEIYVLGLKRRPVSNNRVRERI- 89
 Db 2069 GLYALATHRCYSGLYPHECYSGARASVALTYRASPMETASPGINSERPHEARGILETYR 2128
 Qy 90 SOEVAAGLEFYGNRMW-----LEKPSGFAQPPRLREVAVRKYK 129
 Db 2129 SERVALILIELVSSERTIRILLYSSERGLVALCYSISSERGLYPH-----EVAL----- 2178
 Qy 130 RRRSFYRIFFDSGFTPHGEPBSQRMWLSY-----TANNREYALLRHPBPRLVLCVHGTEMG 187
 Db 2179 ---GLYALALYSCSPHETHVALSERGLG-----LVALPRSRG 2215
 Qy 188 RAPDLAVFRAMKLHDELGLNIWPIVPMH-----GPRCQGLPKGAVPRG-----EDVIDD 238
 Db 2273 LYGLNLIELBALHNSGLYPRASPEERHISVALTYRALAGYASVILBALAARGLEASNA 2332
 Qy 239 VHG-----TAQAVMDIRLLSWIRSOEBSLIGLNGLSIGYIAS--LVAISLEGLACAIL 292
 Db 2273 LYGLNLIELBALHNSGLYPRASPEERHISVALTYRALAGYASVILBALAARGLEASNA 2332
 Qy 293 GAVPAD-LIELGRHCGLRHKOPRRHTVMAEPICRMISPLTFPLVPMGPRFIYAGIAD 351
 Db 2333 SNFSEERLYSMETPEHGLYHISPRQNLSESRHIS-----GLNGLVASPRRILEPHSE 2386
 Qy 352 RL 353
 Db 2387 RL 2388

RESULT 7

US-09-252-991A-17549
 ; Sequence 17549, Application US/09252991A

Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 17549
; LENGTH: 563
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17549

Query Match 4.6%; Score 102; DB 2; Length 563;
Best Local Similarity 23.7%; Pred. No. 0.078;
Matches 87; Conservative 36; Mismatches 106; Indels 138; Gaps 22;

QY 20 PVLGAMKRFMTHTGLYVQSW-----RDYL-GQQPD-----KLPIAR-PTIALAAQAFR- 66
DB 114 PVLGSG-----IQPWPVDSRDPRLADQGDAGRLGVALLPRLALREAGRR 163
QY 67 --DEIVLLGLKARRPVSN-HRVFERISQEVAAGLEFYGNRRRLLEKSGFPAQPPPL----- 119
DB 164 LDDRPRRGGRKRRKVRRAHLL-----GRRIHQPRPAIRPAGPPVAPRT 208
QY 120 ----TEVAVRKYVDKRRRSFYRIFPDGSFTPHRGEBSGRMLSYTANNEVALLRHPEPR 175
DB 209 HQPGVRAALRQLRLDRRAV-----RLGQP-----RPAIHLIGQPEPG 245
QY 176 PWLVCHNCTENGRLPDLAVFRAMKTHDELGLNTVPLP--MGPRQGGPK-----GAV 229
DB 246 GAVACPRRGEPRPPARLD-----PAPRLRGLPQGGPRLAAGAG 288
QY 230 FPGSDVLDVHGTQAQVWDIRLLSWIRSOEBSLIGLNGSLGYSIYASLSEGLAC 289
DB 289 RPA-----LHALAAL-----ARAAMLRAG-----GAGG--AAKRLPAGARRAC 326
QY 290 AILGVPAVDLIELLGRHCGRLHKOPRRHTVKAEBTIGMISPLSTPLVPMGRFTYAGI 349
DB 327 AGSGILLVLL-----AKTH-----PPLRRRRRP-ALSPVVLP--VAGGL 365
QY 350 ADRLVHP 356
DB 366 LPALERP 372

RESULT 8
US-09-902-540-16812
; Sequence 16812, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Steven C.
; APPLICANT: Slater, Gregory J.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(115849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 16812
; LENGTH: 981
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-16812

Query Match 4.5%; Score 99.5; DB 2; Length 981;
Best Local Similarity 21.4%; Pred. No. 0.34;
Matches 66; Conservative 41; Mismatches 105; Indels 97; Gaps 14;

QY 107 EKPSGFFAQPPL-EVAVRKVKRRRSFYRIFPDGSFTPHRGEBSGRMLSYTANREY 165
DB 718 ELPS--VAKPEYTPNVEQKQPER-----FTSLVRPDRFKPKK----- 757
QY 166 ALLRHPEPRPWLVCVHGTENGARPLDLAVFRAMKTH-----DE--LGLNTVPMPLPM 216
DB 758 -----LPVIVEYG-----GPTTVVHKSAAMLMQGVADQGFVLVKIDGRGRL 803
QY 217 HGRPGGLPKGAVPGEVDLDVHGTQAQVWDIRLLSWIRSOEBSLIGLNGSLGYSI 276
DB 804 RGAKTREVXGD-FSGVTLEDOVALQALAKEVEL-----DLRVIGISGWSFGGYM 854
QY 277 ASL-VASLEBGLACAILIGVPADLIE-----LGRHCGLRHKDPRR----- 316
DB 855 SALAVLRKPDVFKAGVAGAPVMDLDVDTYTERYLGVPQSPRAYEKSSLLTYAKODKP 914
QY 317 -----HTVKMAEPICRMISPLSTPLVPMGRFTYAGIADRLVHPRE 358
DB 915 MGKLLIHGTADDVVFHTLTKSDALFRAGKPHDLPRLGLTMM-----VPDPLVTER- 968
QY 359 QVTRLEWHW 367
DB 969 QWERVMDHF 977

RESULT 9
US-09-902-540-13877
; Sequence 13877, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Steven C.
; APPLICANT: Slater, Gregory J.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(115849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 13877
; LENGTH: 748
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-13877

Query Match 4.5%; Score 98.5; DB 2; Length 748;
Best Local Similarity 21.7%; Pred. No. 0.29; Indels 137; Gaps 23;
Matches 100; Conservative 52; Mismatches 171;

QY 40 WRDYLGOQPPKLPARPTIALAAQAFRDEIVLLGLKAR----- 77
DB 84 WKRYLGARPLP--DTLTFEGRAY-----TVGVWMBESFDYRGQLWVRELVPRLPS 136
QY 78 RPVSNNRVFERISOEV--AAGLEFYGNRRRLLEKSGFPAQPPPLTVAVRAKYKDR----- 130
DB 137 RTAHNWRVGVRLADGVLOARVETLTHIARELEAOYG---QDTRMHDIAVEPLQOESLVGR 193
QY 131 -RRSFYRIFPDGSFTPHRGEBSGRMLSYTANNEVALLRHPEPRPWLVCVHGTENGRA 189
DB 194 VRSTLYLL-----AGAAFLLLVAGANTVLLAKRAATRABELAIH-VALGAG 240
QY 190 P-----LDL-----AVFRAMKTHDELGLNTVPMPLPMHGR----- 220
DB 241 PGLAMRFLMESLILSLAGALGVLLAAW-----GVRALLAAEPRLPRVDEGVNATV 294
QY 221 -----GGGL-----PKGAVFGEEDVLDVHGTQAQVWDIRLLSWIRS 258

Db 295 LIFSLGLSLALLGLLTLALRAARPPGALALAGSGRTSGGSART---RAL--VVG 349
Qy 259 QEEESLIGLNGSLGVIASLVASLEEGLA---CALLGV---PVADIELLGHGCLARK 312
Db 350 QALALVLLVGAALLGRSLMGLSLDPGRTRETVAVLSLVPARK--EAQRH---NV 403
Qy 313 DRRHTVK--MAPIGRMISPLSLPLVWMP---GRFIYAGIADRLVHPR--QVTRLWE 365
Db 404 QLOEHLISRLAALPGVRAVAVASSPDEGPGADGFFIVLNRPDEVGNFEDFGLARBE 463
Qy 366 HNGKEPIVWYPGHGTGFPGSRPRRVQALLESGLDLP 405
Db 464 RTGSAAE---YRVASEGYFALGIPLVGRGLFDARDTVDAF 500

RESULT 10
US-09-902-540-15329
; Sequence 15329, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 15329
; LENGTH: 1178
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-15329

Query Match 4.5%; Score 98.5; DB 2; Length 1178;
Best Local Similarity 23.7%; Pred. No. 0.59;
Matches 85; Conservative 36; Mismatches 109; Indels 129; Gaps 20;
Qy 17 LSPVLSGAMGPFMTGLVVAQSWRDYLGQOPDKPIARPTIALAAQAFRDEI----- 69
Db 785 VAPF--GAPMFT-----RDLASVSDALKER-----FQALGHELOAKLRD 825
Qy 70 -----VLGLKARRP---VSNR-----VFRIS--QEVAGLE-----F 99
Db 826 GGPVVTTLALLEGPHRPSRLMTAHLVVDGVSWRVLFEQLDLAVEQLGGEGECKLPRT 885
Qy 100 YGNRRMLEKPGGFPAQPPLTE---VAVRKVKDRRSRFRVIFPDGFTPHGSRGS-Q 153
Db 886 TSVRMLERLQ--YASPRLEBLAWTAQKQBQPPRIVER-----PGLVFTSJE 936
Qy 154 RWLSTANNREYALLLRH--PEBRPMLVCVHTEMGRAPDLAVFPAKMLHDELGINIWP 212
Db 937 RELTALSSSESRLLINLEPE-----AYRCEIGEL--LATLAWTHTWGTGHEVLL 986
Qy 213 VLPWHPGPGQGLPKGAVFPGEDVLDVHGTAQAVMDIRLLSWINS----- 258
Db 987 DIEGHG-----REDFVEFDV-----DLSTRVGLMISLIVPRLTPRVRGS 1025
Qy 259 -----QEEESLIGLNGSLGVIASLVASLEGLACALISGVADLIELLGRHGLRH 311
Db 1026 AATVTVKQRLRSVPNRGL--GIGVLYRISPDALRLARLAPQA---EVLFRYGGQYH 1078

RESULT 11
US-09-602-777A-118
; Sequence 118, Application US/09602777A
; Patent No. 6831165
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus

; APPLICANT: Krogner, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
; FILE REFERENCE: BGI-128CP
; CURRENT APPLICATION NUMBER: US/09/602,777A
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: DE 19931636.8
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19932125.6
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932126.4
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932127.2
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932128.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932129.9
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: DE 19932226.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932920.6
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932924.9
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932928.1
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932930.3
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932933.8
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932935.4
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932973.7
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19933002.6
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19933003.4
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19933005.0
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19933006.9
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19941378.9
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19941379.7
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19941390.8
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19941391.6
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19942088.2
; PRIOR FILING DATE: 1999-09-03
; NUMBER OF SEQ ID NOS: 442
; SEQ ID NO 118
; LENGTH: 668
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-602-777A-118

Query Match 4.5%; Score 98; DB 2; Length 668;
Best Local Similarity 20.8%; Pred. No. 0.27;
Matches 91; Conservative 52; Mismatches 157; Indels 138; Gaps 25;
Qy 41 RDVYGGQDPDKPIARPTIALAAQAF-----RDEIVLGLKARRPVSNR- 84
Db 289 KEFLGGKDFQVPVFTPESTSLQGLATTKNPLVLTLLNNVSTETVTVPL--NDPTTEHH 346

```
QY 85 --VERISOEVAAGLEFYGNRRLEKPSGFFAOPPELLEEV----AVRKVKORRRSFYRIE 138
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 347 IDLEPHVAHVAVATSPDLQDEIIVQAA--FTEAPTLRLAELPGALEAVKAPLQF---- 400
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 139 FDSGTFPHGEGSGQRMWSTYANNREYALLR---HPEPRPWLVCVHGTEKRAELDLAV 195
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 401 -----ENAGGQTRQHWAT--SADGKIPYFITGAEEBQNTLVHAYG----- 441
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 196 FRAWKHLDELGLNIWMPVLPWNGP--RG--QGLPKGAVF-----PGEDVLDDVHGTA---- 243
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 442 -----GEFVSLTPSHSPTRGLAMLEKGYTFVEANLRGGGEPERHNSQATKLN 489
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 244 -QAVW-DIRRLSLWI--RSQEEESLIGLNGSLGGYIAS-LVASDEGLACAILGVPAVD 298
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 490 RMKVEDHRAVLADLVERGYATPEQIAIRGSGNGGLTSGALTQYPEAFGAALVQVPLAD 549
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 299 LIEELGRHCGLRHKDPRRHTTK-----MAE-----PIGRMI-----SPLSLTPVLP--M 340
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 550 ML-----RYHTWSAGASWMAEYGNPDPPEERAVIEQYSPQAVVGEKRIY 595
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 341 PGRFYAGIADRLVHPREQVRLW-----EHMGKPEIWMVYPGGHTG-----F 382
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 596 PPALVTTSTRDRVPAH--ARLFPAQLLDAGQAVDYENTE-----GSHAGADNKKQTA 648
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 383 FQSRPVRREFVQAALBQSG 400
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 649 FVESLIYTWIEKTLDDQG 666
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

```
RESULT 12
US-09-105-537-22
; Sequence 22, Application US/09105537A
; Patent No. 6265202
```

```
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.438US1
; CURRENT APPLICATION NUMBER: US/09/105,537A
; CURRENT FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 22
; LENGTH: 402
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
US-09-105-537-22
```

```
Query Match 4.4%; Score 97; DB 2; Length 402;
Best Local Similarity 22.8%; Pred. No. 0.16;
Matches 75; Conservative 35; Mismatches 139; Indels 80; Gaps 15;
```

```
QY 110 SGFFAOPPELLEVAVKVDRRRSFYRIFFDSGTFPHGEGSGQRMWSTYANNREYALL 169
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 6 TGAULTQRP--LGRYTRAVADRE-----LGTILLETRGIHWI--HAANGDPYATVL 51
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 170 RHPEPRPWLVCVHGTEKRAELDLAVFRWKHLDE--LGNIYMPV--LPMHGRGQGLPKG 227
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 52 RQADDPY--PAYEVRARAGALSPTSQSWTADHALAASILCSTDGVSQADGVVPVQ 109
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 228 AVFPGEDVLDDVHGTAQAVMDI-----RRLSWIRSOEESLIGLNGSL-----LGYYIA 277
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 110 VLSYEGCGRLERQVLPAGADVPEGGQRAVVEGIRHETIEGLAPRPSASYAFELLGGFVR 169
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 278 SLVASLEBGLACAILGVPV-----ADLIELGRHCGLRHKDPRRHTTKMAEPI----- 325
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 170 PAVTA-----AAAAYLGVADRRADPADLERLRLPSDSLAPQSIRTVAAAGALAEELTA 225
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 326 -----GRMISPLSLTPVPMGRFYAGIADRLVHPREQVTRLMENHGKPEIWMY 376
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

```
DB 226 LIAQSDSPGALLSLKGTAAVOLTGNANVA-----LLANP-----EQW--RELCORP 271
QY 377 GGHTEGFPQSRPVRREFVQAALBQSGLLDAP 405
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 272 G-----LAAAAYBETRLRYDPP 287
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

```
RESULT 13
US-09-105-537-4
; Sequence 4, Application US/09105537A
; Patent No. 6265202
```

```
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.438US1
; CURRENT APPLICATION NUMBER: US/09/105,537A
; CURRENT FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 3782
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
US-09-105-537-4
```

```
Query Match 4.4%; Score 97; DB 2; Length 3782;
Best Local Similarity 22.8%; Pred. No. 5.3;
Matches 75; Conservative 35; Mismatches 139; Indels 80; Gaps 15;
```

```
QY 110 SGFFAOPPELLEVAVKVDRRRSFYRIFFDSGTFPHGEGSGQRMWSTYANNREYALL 169
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 6 TGAULTQRP--LGRYTRAVADRE-----LGTILLETRGIHWI--HAANGDPYATVL 51
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 170 RHPEPRPWLVCVHGTEKRAELDLAVFRWKHLDE--LGNIYMPV--LPMHGRGQGLPKG 227
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 52 RQADDPY--PAYEVRARAGALSPTSQSWTADHALAASILCSTDGVSQADGVVPVQ 109
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 228 AVFPGEDVLDDVHGTAQAVMDI-----RRLSWIRSOEESLIGLNGSL-----LGYYIA 277
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 110 VLSYEGCGRLERQVLPAGADVPEGGQRAVVEGIRHETIEGLAPRPSASYAFELLGGFVR 169
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 278 SLVASLEBGLACAILGVPV-----ADLIELGRHCGLRHKDPRRHTTKMAEPI----- 325
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 170 PAVTA-----AAAAYLGVADRRADPADLERLRLPSDSLAPQSIRTVAAAGALAEELTA 225
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 326 -----GRMISPLSLTPVPMGRFYAGIADRLVHPREQVTRLMENHGKPEIWMY 376
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 226 LIAQSDSPGALLSLKGTAAVOLTGNANVA-----LLANP-----EQW--RELCORP 271
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 377 GGHTEGFPQSRPVRREFVQAALBQSGLLDAP 405
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 272 G-----LAAAAYBETRLRYDPP 287
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

```
RESULT 14
US-09-902-540-10121
```

```
; Sequence 10121, Application US/09902540
```

```
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(115849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
```

SEQ ID NO 10121
LENGTH: 596
TYPE: PRT
ORGANISM: Myxococcus xanthus
US-09-902-540-10121

Query Match 4.4%; Score 96.5; DB 2; Length 596;
Best Local Similarity 23.7%; Pred. No. 0.33;
Matches 97; Conservative 38; Mismatches 140; Indels 135; Gaps 23;

QY 8 GTHERSAFPLSPVLSGANGFPMHTGLVYASWRDY---LGOQPDILPIAR----- 55
DB 169 GTHKS-----LGAAPSLIRREHGHLMKRDRIIGRELDLGVKRYGNGPQ 220
QY 56 ---PTALAA---QAFDEIVLLGLKARPVSNHRVFERISQEVAAEGFYGNRRMLE- 107
DB 221 VPPVSTIGMAAMLOTATRRPALLAKAE---CARGFQVSVR-----NRPVVA 268
QY 108 -KPSGFPAQ---PPLTEVAVKVKDRRSFYRIIPDSGFTPHPEPSQRMWLYTANN 162
DB 269 PFDGQLLPERVAPYFEEYAVPDARQAARFVHEFV-----EP----- 307
QY 163 REVALLR---HEPRPMLVCV---HGTMGRAPLDLA-----VPRAMKLDDEL 206
DB 308 -----LRDAVRAEPFPYVACLVADGRGRLAELAROPDGHAAHQISRLADFTROA 361
QY 207 LNIIV---MPVLPMHGPRGGLPKGAVFGEDEVLD-----DVHGTAAVMDIRRLSMIR 257
DB 362 KTIIVEVHGVLVYAG-----GDDILAFVTPMDALACQA-----LAVAFR 402
QY 258 SOEESLIGLN-----GLSLGGYIASIVASLEGLAC--AIIIGVPVADLITELLGRHC 307
DB 403 ATLEALAGTRAEVPTLVGIGVLESIGELLEGRABRAAKDAGRDAVLAVAKHA 462
QY 308 GLRHKDPRRHTVMA-EPICRMISPLSL--TPLVPMGRFTY--AGIADR 352
DB 463 GRB---RLMTAPWATDPVERLTRDMALLSSDTQPLKKVHEVVALARR 508

RESULT 15

US-09-902-540-15882
Sequence 15882, Application US/09902540
Patent No. 6833447
GENERAL INFORMATION:
APPLICANT: Goldman, Barry S.
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
FILE REFERENCE: 38-10(15849)B
CURRENT APPLICATION NUMBER: US/09/902,540
PRIOR FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: 60/217,883
PRIOR FILING DATE: 2000-07-10
NUMBER OF SEQ ID NOS: 16825
SEQ ID NO 15882
LENGTH: 416
TYPE: PRT
ORGANISM: Myxococcus xanthus
US-09-902-540-15882

Query Match 4.4%; Score 96; DB 2; Length 416;
Best Local Similarity 24.1%; Pred. No. 0.21;
Matches 76; Conservative 35; Mismatches 115; Indels 90; Gaps 18;

QY 8 GTHERSAFPLSPVLSG---AMGP---FMHT-----GLVYA-----QSWRDYLGQ 47
DB 76 GEVDTSADARALPYVLSGCGVLPISGAFIRTDATAGTPVGLANVSLDQGRQILTLL 135
QY 48 PDLPL--IARPT--ALAAQAFRDE-----IVLLGLKARRPVSNHRVFERISQEVAAEGF 99
DB 136 ADALKGVABPAAALPVDPTIDEARIRAILLAIST--AGNKGVLRRTVRLVLER 192

QY 100 YG---NRRW-LEKPSGFPAQPPPLTEVAVRKVKDRRSFYRIIPDSGFTPHPEGSGSR 154
DB 193 YNAETEOLEMAHEDPYTDMGEAPYDIDVI-----GHNSAYMRLEA-----GTPGGQR 240
QY 155 WLSYANNREVALLLRH-----PEPRPMLVCVHGT-----EMGRAPLDLA---VFRAM 199
DB 241 ---FTSPPLDVLFRIRHRRHRSPPABEGVVTFFHPLMRQEGEMERAVLDLSFGSMCIRC 297
QY 200 KLHDELGLNIYMPVLPMHGPRGGLPKGAVFGEDEVLDVHGTAAVMDIRRLSMIRSQ 259
DB 298 REPEDLFPGLPPLLELHTPDQSI-----SLRGEIRYVTS 334
QY 260 EESLIGLNLGLSGY 275
DB 335 RADGTV-LCGLSVSPY 349

Search completed: March 23, 2006, 05:30:19
Job time : 29.268 sec

This Page Blank (uspte

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: March 23, 2006, 06:14:33 ; Search time 86.8513 Seconds
(without alignments)
1986.885 Million cell updates/sec

Title: US-10-617-038-29

Perfect score: 2195 1 MASSASDGTHERSAPRLSP.....AALQSGLLDAPRTQDRSA 413

Sequence: BLOSUM62

Scoring table: Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

- 1: Published Applications_AA_Main:*
- 2: /cgn2_6/pcodata/1/pubppaa/US07_PUBCOMB.pep:*
- 3: /cgn2_6/pcodata/1/pubppaa/US08_PUBCOMB.pep:*
- 4: /cgn2_6/pcodata/1/pubppaa/US09_PUBCOMB.pep:*
- 5: /cgn2_6/pcodata/1/pubppaa/US10_PUBCOMB.pep:*
- 6: /cgn2_6/pcodata/1/pubppaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2195	100.0	413	4 US-10-617-038-29	Sequence 29, Appl
2	146.5	6.7	366	4 US-10-425-115-286397	Sequence 286397,
3	146.5	6.7	369	4 US-10-425-114-39123	Sequence 39123, A
4	146.5	6.7	369	4 US-10-425-114-47261	Sequence 47261, A
5	130	5.9	303	4 US-10-424-599-196342	Sequence 196342,
6	121	5.5	328	6 US-11-097-143-27990	Sequence 27990, A
7	118.5	5.4	322	5 US-10-739-930-9602	Sequence 9602, Ap
8	112	5.1	574	4 US-10-282-122A-62026	Sequence 62026, A
9	111.5	5.1	311	4 US-10-437-963-183563	Sequence 183563,
10	110	5.0	481	4 US-10-424-599-227907	Sequence 227907,
11	107.5	4.9	1470	4 US-10-425-115-206590	Sequence 206590,
12	106.5	4.9	691	3 US-09-976-674-43	Sequence 43, Appl
13	106.5	4.9	691	5 US-10-982-512-43	Sequence 43, Appl
14	106	4.8	452	4 US-10-425-114-60109	Sequence 60109, A
15	104.5	4.8	742	4 US-10-369-493-15631	Sequence 15631, A
16	104.5	4.8	742	4 US-10-369-493-16006	Sequence 16006, A
17	104.5	4.8	742	4 US-10-369-493-16360	Sequence 16360, A
18	103	4.7	810	4 US-10-425-115-274900	Sequence 274900,
19	102.5	4.7	980	5 US-10-794-514A-342	Sequence 342, App
20	100.5	4.6	586	5 US-10-755-415-241	Sequence 241, App
21	100	4.6	740	4 US-10-437-963-115057	Sequence 115057,
22	99.5	4.5	424	4 US-10-425-114-64299	Sequence 64299, A
23	98.5	4.5	1415	4 US-10-425-115-206589	Sequence 206589,
24	98	4.5	296	4 US-10-425-115-224377	Sequence 224377,
25	98	4.5	668	5 US-10-721-922A-178	Sequence 178, App
26	98	4.5	706	3 US-09-738-626-3879	Sequence 3879, Ap
27	98	4.5	706	5 US-10-494-541-22	Sequence 22, Appl

28	97	4.4	402	3 US-09-861-289-22	Sequence 22, Appl
29	97	4.4	402	3 US-09-860-846-22	Sequence 22, Appl
30	97	4.4	402	3 US-09-988-384B-22	Sequence 22, Appl
31	97	4.4	402	3 US-09-836-821-22	Sequence 22, Appl
32	97	4.4	402	4 US-10-271-889-22	Sequence 22, Appl
33	97	4.4	402	4 US-10-398-605-22	Sequence 22, Appl
34	97	4.4	706	4 US-10-425-115-309560	Sequence 309560,
35	97	4.4	3782	3 US-09-861-289-4	Sequence 4, Appl1
36	97	4.4	3782	3 US-09-860-846-4	Sequence 4, Appl1
37	97	4.4	3782	3 US-09-988-384B-4	Sequence 4, Appl1
38	97	4.4	3782	3 US-09-836-821-4	Sequence 4, Appl1
39	97	4.4	3782	4 US-10-398-605-4	Sequence 4, Appl1
40	97	4.4	1595	4 US-10-425-115-238559	Sequence 238559,
41	96.5	4.4	421	4 US-10-092-900A-88	Sequence 88, Appl
42	96	4.4	442	4 US-10-425-115-223184	Sequence 223184,
43	96	4.4	711	4 US-10-425-114-46399	Sequence 46399, A
44	96	4.4	809	6 US-11-097-143-4965	Sequence 4965, Ap
45	96	4.4			

ALIGNMENTS

RESULT 1					
US-10-617-038-29					
; Sequence 29, Application US/10617038					
; Publication No. US20040057963A1					
; GENERAL INFORMATION:					
; APPLICANT: Andersen, Peter					
; APPLICANT: Rosenkrands, Ida					
; APPLICANT: Strjhn, Anette					
; TITLE OF INVENTION: Therapeutic TB Vaccine					
; FILE REFERENCE: SSIUSAUSA					
; CURRENT APPLICATION NUMBER: US/10/617,038					
; CURRENT FILING DATE: 2003-07-11					
; PRIOR APPLICATION NUMBER: DK PA 2002 01098					
; PRIOR FILING DATE: 2002-07-13					
; PRIOR APPLICATION NUMBER: US 60/401,725					
; PRIOR FILING DATE: 2002-08-07					
; NUMBER OF SEQ ID NOS: 187					
; SOFTWARE: PatentIn version 3.2					
; SEQ ID NO 29					
; LENGTH: 413					
; TYPE: PRT					
; ORGANISM: Mycobacterium tuberculosis					
US-10-617-038-29					
Query Match					
Best Local Similarity 100.0%; Score 2195; DB 4; Length 413;					
Matches 413; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
QY	1	MASSASDGTHERSAPRLSPVSGAMGPMHTGLVVAOSMRDYLGOQPKLPIARPTL	60		
DB	1	MASSASDGTHERSAPRLSPVSGAMGPMHTGLVVAOSMRDYLGOQPKLPIARPTL	60		
QY	61	AAQARDEIVLGLKARBPVSNHRVFERRISQVAAGLEFGYGRNMLEKSGFFAQPPLYT	120		
DB	61	AAQARDEIVLGLKARBPVSNHRVFERRISQVAAGLEFGYGRNMLEKSGFFAQPPLYT	120		
QY	121	EVAVRKVKDKRRSPFRIFPDGFTPHPGSPGQRLSTYANNREYALLLRHBPWPVLC	180		
DB	121	EVAVRKVKDKRRSPFRIFPDGFTPHPGSPGQRLSTYANNREYALLLRHBPWPVLC	180		
QY	181	VHGTMGRAPLDLAFAFMKLDHDELGLNIVMVLPMHGRGGLPKGAFFPSEDVLDVH	240		
DB	181	VHGTMGRAPLDLAFAFMKLDHDELGLNIVMVLPMHGRGGLPKGAFFPSEDVLDVH	240		
QY	241	GTAQAVMDIRLLSWIRSOEBSLIGLNGLSIGYIASLVASLEGLACALIGVAVDLI	300		
DB	241	GTAQAVMDIRLLSWIRSOEBSLIGLNGLSIGYIASLVASLEGLACALIGVAVDLI	300		
QY	301	ELLGHHGGLRKDPRRHVVKAEPFGMTSPSLSTPLVPMGRFYAGIADRLVHPRQV	360		
DB	301	ELLGHHGGLRKDPRRHVVKAEPFGMTSPSLSTPLVPMGRFYAGIADRLVHPRQV	360		

Db 301 ELGRRICGLRHQDPRRHTVYMAEPIGRMTSPSLTPRYVPMGRFTYAGIDRLVHPREQV 360

Qy 361 TRLMENHKREIIVYRPGHTGFFQSPRPVRRFYQALLESGLDAPRTQRDRA 413

Db 361 TRLMENHKREIIVYRPGHTGFFQSPRPVRRFYQALLESGLDAPRTQRDRA 413

```

RESULT 2
US-10-425-115-286397
; Sequence 286397, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 286397
; LENGTH: 366
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_24291C.1.pep
US-10-425-115-286397

```

```

RESULT 3
US-10-425-114-39123
/ Sequence 39123, Application US/10425114
/ Publication NO. US2004003488A1
/ GENERAL INFORMATION:
/ APPLICANT: Liu, Jingdong
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Screen, Steven E
/ APPLICANT: Tabaska, Jack E
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
/ FILE REFERENCE: 38-21(5313)B
/ CURRENT APPLICATION NUMBER: US/10/425,114
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 73128
/ SEQ ID NO 39123
/ LENGTH: 369
/ TYPE: prt
/ ORGANISM: Zea mays
/ FEATURE:

```

```

; OTHER INFORMATION: Clone ID: 700212728_F1.1.pep
US-10-425-114-39123

Query Match      6.7%; Score 146.5; DB 4; Length 369;
Best Local Similarity 24.6%; Pred. No. 2.4e+05;
Matches 59; Conservative 31; Mismatches 81; Indels 69; Gaps 8

```

```

RESULT 4
US-10-425-114-47261
; Sequence 47261, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 47261
; LENGTH: 369
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700266090_f1.1.pep
US-10-425-114-47261

Query Match      6.7%; Score 146.5; DB 4; Length 369;
Best Local Similarity 24.6%; Pred. No. 2.4e-05;
Matches 59; Conservative 31; Mismatches 81; Indels 69; Gaps 8;

QY      208 NTVMPVL--PMHGPRGGGLPKGA-----VFPGEVDLDVHGTAQAAMDRLRLSMIR 257
           ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      149 NATATWLSPPYGGQRRPMSORGAKKQCVSDLLLGKATIDEA-----RSLYMQ 198

QY      258 SOEESILGLNGLSGGYIASLVASLE-----EGV----- 287
           ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      199 NENGYGKMGICGLSMGCVHAAMGSLNHTPVATLPLAHSVAVPPCEGYYKATAMDAL 258
           ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      288 --ACALGVPAVDLIELGRHCGLRHKDPRRHTVKMAEPIGRMISPLSLT---PLVPM 340
           ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      259 RKDAVLTQDVTLLLEDAAKSGSITIEQV-----DRLRSVLSLTDVTRFVPKN 308

QY      341 PGRFTYAGIADLVLPREGQVTRLMENHCKRPETIVYPGGHTG--FQSGRPARRVQAALRG 398
           ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      309 PQAVIFVGATDDGYIPRHSVMELOKAMPGSEVRWVTGHWSSFFLHNDAPFRKAIVDALDR 368

RESULT 5
US-10-424-599-196342
; Sequence 196342, Application US/10424599

```

```

; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 195342
; LENGTH: 303
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(303)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3647_19323C.1.pep
; US-10-424-599-196342

```

Query Match	5.9%	Score 130;	DB 4;	Length 303;
Best Local Similarity	29.1%;	Pred. NO. 0.00077;		
Matches 66; Conservative	23;	Mismatches 88;	Indels 50;	Gaps 12

```

Oy 203 BELGANTIVPVLPMHGPRQGLPKGVN--PGESEVLDVHG-----TAQAAMDRLRLLSWI 256
Db 88 EEQLGLALPVL--FTYARAPVLPRGXKRCVSDLL--LGRPLYETNP-----RPLHW 137
Oy 257 RSQEESLGLNGLSGTYIASIVASLEBGLACALIGVPVADLIELGRHC----- 308
Db 138 CLDAGFSKMGKGLGSLNCGVHAAMVGSLLP-----TPATL--PFLSPSAVAFCEGI 188
Oy 309 LRH-----KOPRRHTYKMA--EITGMISPLSLT-----PLVMPGCFITYAGIADRL 353
Db 189 LKHGTAMEALRLDLAAQKAMVLTLEBRREMRNVLTSTDVTRPEI PKPNPAVIEVAATDDG 248
Oy 354 VHPRQVYTRLMHMGKPELIVYTPGCHTGF--QSRVVRPVOALALRQ 398
Db 249 YIPKISVIELQKAPGSEVWRMTGGVASSFILHNDEFRAIKQGLRR 295

```

RESULT 6
US-11-097-143-27990
Sequence 27990, Application US/11097143
Publication No. US20050208558A1
GENERAL INFORMATION:
APPLICANT: Venter, J. Craig
APPLICANT: et al.
TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
FILE REFERENCE: CL000728
CURRENT APPLICATION NUMBER: US/11/097,143
CURRENT FILING DATE: 2005-04-04
PRIOR APPLICATION NUMBER: 60/157,832
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: 60/166,191
PRIOR FILING DATE: 1999-10-19
PRIOR APPLICATION NUMBER: 60/161,932
PRIOR FILING DATE: 1999-10-28
PRIOR APPLICATION NUMBER: 60/164,769
PRIOR FILING DATE: 1999-11-12
PRIOR APPLICATION NUMBER: 60/173,383
PRIOR FILING DATE: 1999-12-28
PRIOR APPLICATION NUMBER: 60/175,693
PRIOR FILING DATE: 2000-01-12
PRIOR APPLICATION NUMBER: 60/184,831
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: 60/191,637

```

? PRIOR FILING DATE: 2000-03-23
? NUMBER OF SEQ ID NOS: 43008
? SOFTWARE: FastSeq for Windows Version 4.0.
? SEQ ID NO 27990
? LENGTH: 338
? TYPE: PRN
? ORGANISM: DROSOPHILA
US-11-097-143-27990

```

Query Match	Score	DB	Length
5.5%	121	6	338

Matches 68; Conservative 41; Mismatches 84; Indels 84; Gaps 16.

165 YALLLRHPEPR---PWLVCVHGT--EMGRAPLDLAVFRAWKLDDELGLNIWMPVLPMHG 218

Db 94 HAFWVTQPEERSKSSPTLLYFHGNAGNMGH-----RMQNVWGIYHLLHCNVLMEVE----- 144

219 PRGGGLPKGAVFPGEDVLDDVHGTAAQAVWDIRLLSWIRSQE--EESLIGLNGLSLGGYI 276

DB 145 -RGYGLSIGV--PIERGL-----VIDAKAALDIHTIKHLDLHSQLLLFGRSLGGAV 192

[illegible]

DOCUMENT INFORMATION

Db 238 FKIKYHSMSKIGKCSVPDLFTSGIADNI.VPPRMBRALYTKCGSEIKRLLE-----F 288

QY 376 PGCH-----TGFFQSRPVRRFVQALQSGLLDAP 405

Db 289 PGGSHNDTWIVDGYQA--IGGFL-AELQQQLKAP 322

RESULT 7

; Sequence 9602, Application US/10739930

ADDITIONAL INFORMATION: David R

TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH PLANTS AND USES THEREOF FOR PLANT IMPROVEMENT

CURRENT APPLICATION NUMBER: US/10/739,930

NUMBER OF SEQ ID NOS: 11088

; LENGTH: 322
TYPE: DPT

```

; ORGANISM: ILICUM Aescivum
; FEATURE:

```

LOCATION: (1) .. (322)

FEATURE:

US-10-739-930-9602

Query Match	3.4%	Score 118.5	DB 5	Length 322
Best Local Similarity	25.3%	Pred. No. 0.011		

	0	1	2	3	4	5	6	7	8	9
0	0	1	4	9	16	25	36	49	64	81
1	1	3	8	15	24	35	48	63	80	99
2	4	8	16	25	36	49	64	81	100	121
3	9	15	25	36	49	64	81	100	121	144
4	16	24	36	49	64	81	100	121	144	169
5	25	35	49	64	81	100	121	144	169	196
6	36	48	64	81	100	121	144	169	196	225
7	49	63	81	100	121	144	169	196	225	256
8	64	80	100	121	144	169	196	225	256	289
9	81	99	121	144	169	196	225	256	289	324

[illegible]

OR 258 SOREESLTGNGSLGGYASLVASLEEGIACATLGP-----VADLIELGRHC--- 307

Db 180 AEPGYKMGICGLSMGGVHAAMAGSL--LSTPIPTLPFLAPHSAVVPFCEGLYRHATAW 236

QY 308 -GLRH-----KDPKRHTVKMA-----EPIGRMISPLSLTPLVPMGRFIYAGIADRLVH 355

Db 237 EALREDAALAKDATSLTEDAASGITIEQVXRQIT-----DDGYI 276

Db 202 RHATANDALRKAAATLAQDVTSLTEDMAQSGTITTEQVBERLSVSLDVTFFPVKXP 261
QY 342 GREPIYAGIADRLVHPREOVTRLMEHWGKEPIW 374
Db 262 QAVIFVGATDDGVIIPRHSIMELOKAW--PGESEW 232

RESULT 10

US-10-424-599-227907
Sequence 227907, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa Thomas J
APPLICANT: Kovalic David K
APPLICANT: Zhou Yihua
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53223)B
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 227907
LENGTH: 481
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_47828C.1.pep
US-10-424-599-227907

Query Match 5.0%; Score 110; DB 4; Length 481;
Best Local Similarity 23.4%; Pred. No. 0.14;
Matches 63; Conservative 38; Mismatches 126; Indels 42; Gaps 8;

QY 161 NNRREYLLLRHPRRW-----LVCV---HGTMGARPDLDVAFRMAKHLDELGLNTWP 212
Db 44 NTRGYTLKSHYLPSPFPDTSLPVYICHGNSGCADANEAAIL-----LPSNITVF 97
QY 213 VTPMHGPRGQGLPKGAVFPGEVDLDVHGTAQAMDIRLLSWIRSGEESLIGLGLSL 272
Db 98 TIDF---SSGSLSDG-----DIVSLGMEKDLKAVSVTLRSNKQISRLGLMGRSM 145
QY 273 GGIYASLVASLEBGLACILGVAVADLIELGRHCGLRHDPRRHTVKAAPRIGMI--- 329
Db 146 GAVTSLYGABDPSIAGVNLDSAFSNLYDLMELADVYKIRLPKFVYKAAVQYMRVIEK 205
QY 330 -----SPLSLTPLVMPGRFTIAGIADRLVHPREOVTRLMEHWGKEPIWYFGCHT 380
Db 206 KAKFDIMNINCLQVAKTFIPVLFHAS--DDKFIQPHSDLSI EAYAGDKVYIKFDGDHN 264
QY 381 GFGQSRPVARFVQAALEGSLDAPRTOR 409
Db 265 ---SSRPQFFYDSVSIFFTNVHLHPVPR 290

RESULT 11

US-10-425-115-206590
Sequence 206590, Application US/10425115
Publication No. US20040214272A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53222)B
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 369326
SEQ ID NO 206590
LENGTH: 1470
TYPE: PRT

ORGANISM: Zea mays
NAME/KEY: unsure
LOCATION: (1)..(1470)
OTHER INFORMATION: unsure at all Xaa locations
FEATURE:
OTHER INFORMATION: Clone ID: MRT4577_119997C.1.pep
US-10-425-115-206590

Query Match 4.9%; Score 107.5; DB 4; Length 1470;
Best Local Similarity 20.0%; Pred. No. 1.2;
Matches 84; Conservative 53; Mismatches 135; Indels 149; Gaps 25;

QY 2 AASASDTHERSAFRLSPVLSGA-----MGPFMTG-----LYVAGSWRD 42
Db 4 STTAATGTIAPAIG--SEPAHGAPYITIGLP--LHAASPLPHRTSPGFPAPASPTWGH 61
QY 43 YIGQPDKLPIDRPIT-----ALAAQARDELIVLGLAKRRPVSNHRVFERISQEV 93
Db 62 RVA--PSTIPPTDPLVSTIASIOAVMSRBRERASIALERERALG-----A 108
QY 94 AAGLEFYGNRWLEKPSGFPAPPP-----LTVAVRKVDORRSFYRI 137
Db 109 ALATQWATTQRLIGHPTLADAEPPEDPHASDIDADLIAALHQAQGLHNI---RALYSV 164
QY 138 FPDGFTPHGPGSGQRLSYT--ANNREYAL---LIRH--DEPPRWLVCHTEKGRAP 190
Db 165 VLD-----PASPHYSRMRGQVLLTLRRFVLDHVLIDHTSPRSW--CL----- 207
QY 191 LDIAVFRAMKHLDELGLNTVMPVLPMHGPRGQGLKRGAVFPGEVDLDVHGTAQAVW--- 247
Db 208 MD--SVYLSW--LHDTITVEL-----QDIIRQADADRAQAMIAL 242
QY 248 -----DIRRLSWIRSGEESLIGLGLSGLGIYASLVASLEGLACAI--LGVPA 297
Db 243 EDQFGGNRDARL---HLDQFHLPSQEGLSVGECROM---KGLADSLRDGEFVA 293
QY 298 DLIELGRHCGLRHDP---RHTVKAAPRIGR-----MISPLSLTPLVMPGRPI 345
Db 294 DRTIVLNL---LRGLSPRHGLKALIKRSVPPTTHVYRNELLBELTWANABPFPASAL 350
QY 346 Y 346
Db 351 Y 351

RESULT 12

US-09-976-674-43
Sequence 43, Application US/09976674
Patent No. US20020115843A1
GENERAL INFORMATION:
APPLICANT: Qi, Steve
APPLICANT: Akinsanya, Karen
APPLICANT: Riviere, Pierre
APPLICANT: Junien, Jean-Louis
TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
FILE REFERENCE: 70669
CURRENT FILING DATE: 2001-10-12
PRIOR APPLICATION NUMBER: US 60/240,117
NUMBER OF SEQ ID NOS: 61
SOFTWARE: PatentIn version 3.1
SEQ ID NO 43
LENGTH: 691
TYPE: PRT
ORGANISM: Homo sapiens
US-09-976-674-43

Query Match 4.9%; Score 106.5; DB 3; Length 691;
Best Local Similarity 18.2%; Pred. No. 0.51;
Matches 63; Conservative 63; Mismatches 106; Indels 115; Gaps 14;

QY 9 THERSAFRLSPVLSGAMGPFMTGLVVAQSWRDYLGQOPDKLPIARPTIALAAQAFRDE 68
DB 288 SBDGSKFETVVPKQGRGEFHHIAMFLIQSKEQI-----TVRHLTSGNMEV 335
QY 69 IYLLGKARPVSNHHRFERSIOEVAAGLEFYGNRMLEKSGFPAQPPLEVAVRKYK 128
DB 336 IKIL-----AYDETTOKISASTEGILNRQC-----ISCNFMK 367
QY 129 DRRRSFYRIFPDGGFTPHRPGSQRMWLSY-----TANNREYALI----- 168
DB 368 EQ-----CTYFDASFSPM-----NQHFLLFCGEPRPVVSLSHSTDPNPKYFILESNSMK 417
QY 169 ---LHREPRPMLVCVCHGTEMGRAPLDLAVFRAMKHLDELGLNIVPVLPMHGPRGQGLP 225
DB 418 EAILKKKIKGPEIKILHIDY-ELPLQSLSPKDFMDRNOYALLIMD---EEPGQLVY 472
QY 226 ---KGAVPGEEDVLDDVH---GTAQAVMDIRRL-----L 253
DB 473 DKFHIDMSVLI DMNVIVARPDGRSGFGQGLKILQEIHRRLGSVGVKQITRAVKELKL 532
QY 254 SWIRSQEBSLIGLNGLSLGGYIASLVASLEBGL-ACAILGVPAVDL 299
DB 533 PYIDSKR-----LSIFGKGYGVYIASMLKSDEKLFCGGSVVAPIITDL 575

RESULT 13
US-10-982-512-43
Sequence 43, Application US/10982512
Publication No. US20050059081A1
GENERAL INFORMATION:
APPLICANT: Qi, Steve
APPLICANT: Akinaanya, Karen
APPLICANT: Riviere, Pierre
APPLICANT: Junten, Jean-Louis
TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPV1
FILE REFERENCE: 70669
CURRENT APPLICATION NUMBER: US/10/982,512
CURRENT FILING DATE: 2004-11-05
PRIOR APPLICATION NUMBER: US/09/976,674
PRIOR FILING DATE: 2001-10-12
PRIOR APPLICATION NUMBER: US 60/240,117
PRIOR FILING DATE: 2000-10-12
NUMBER OF SEQ ID NOS: 61
SOFTWARE: Patent version 3.1
SEQ ID NO 43
LENGTH: 691
TYPE: PRT
ORGANISM: Homo sapiens
US-10-982-512-43

Query Match 4.9%; Score 106.5; DB 5; Length 691;
Best Local Similarity 18.2%; Pred. No. 0.51;
Matches 63; Conservative 63; Mismatches 106; Indels 115; Gaps 14;
QY 9 THERSAFRLSPVLSGAMGPFMTGLVVAQSWRDYLGQOPDKLPIARPTIALAAQAFRDE 68
DB 288 SBDGSKFETVVPKQGRGEFHHIAMFLIQSKEQI-----TVRHLTSGNMEV 335
QY 69 IYLLGKARPVSNHHRFERSIOEVAAGLEFYGNRMLEKSGFPAQPPLEVAVRKYK 128
DB 336 IKIL-----AYDETTOKISASTEGILNRQC-----ISCNFMK 367
QY 129 DRRRSFYRIFPDGGFTPHRPGSQRMWLSY-----TANNREYALI----- 168
DB 368 EQ-----CTYFDASFSPM-----NQHFLLFCGEPRPVVSLSHSTDPNPKYFILESNSMK 417
QY 169 ---LHREPRPMLVCVCHGTEMGRAPLDLAVFRAMKHLDELGLNIVPVLPMHGPRGQGLP 225
DB 418 EAILKKKIKGPEIKILHIDY-ELPLQSLSPKDFMDRNOYALLIMD---EEPGQLVY 472
QY 226 ---KGAVPGEEDVLDDVH---GTAQAVMDIRRL-----L 253
DB 473 DKFHIDMSVLI DMNVIVARPDGRSGFGQGLKILQEIHRRLGSVGVKQITRAVKELKL 532

QY 254 SWIRSQEBSLIGLNGLSLGGYIASLVASLEBGL-ACAILGVPAVDL 299
DB 533 PYIDSKR-----LSIFGKGYGVYIASMLKSDEKLFCGGSVVAPIITDL 575

RESULT 14
US-10-425-114-60109

Sequence 60109, Application US/10425114
Publication No. US20040034888A1
GENERAL INFORMATION:
APPLICANT: Liu, Jindong
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E
APPLICANT: Tabaska, Jack E
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(5313)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 60109
LENGTH: 452
TYPE: PRT
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: LIB3632-038-B12_Flt.pep
US-10-425-114-60109

Query Match 4.8%; Score 106; DB 4; Length 452;
Best Local Similarity 24.1%; Pred. No. 0.31;
Matches 86; Conservative 48; Mismatches 149; Indels 74; Gaps 19;

QY 52 PIARPTIALAAQAFRDEIVLLGLKARPVSNHHRFERSIOEVAAGLEFYGNRMLEKSG 111
DB 50 PLVRSPSGAREBA-----PISHPM-GAASVVAARFAFP-----PTPS 89
QY 112 FFAQPPPLETEVAVRKYDRRRSFYRIFPDGGFTPHRPGSQRMWLSYANNREYALLRH 171
DB 90 YGVBRPPSPAAADTE-----VVELSGVPVSRGSGVBARLPTKRGEVWSMYRQ 141
QY 172 PEPRPMLVCVCHGTEMGRAPLDLAVFRAMKHLDELGLNIVPVLPMHGPRGQGLKAVFP 231
DB 142 PGARLTLTSHGN-----AADLG-QMYELFVELSAHNVLMG-YDYSYGQSGSK--P 191
QY 232 GEDVLDVHGTAQAVMDIRRLLSWIRSQEBSLIGLNGLSLGGYIASLVASLEGLACAI 291
DB 192 SE---QNTYADIEAVY-RCLLESYASSEN--ILLGQSVGSGPTLDLASRLPHLRV 244
QY 292 LGVPVADLIELGHRGHLHNDPRRHVTKMAEPIGRMISPLSLPLVPMGPFYAGIAD 351
DB 245 LHSPISS-----GLRWVVPVGHYWF-----DIKNDIKPIVRCPLVH-GRAD 289
QY 352 RLV---HREQVTRLWE-HMKPEIIVYDGHGTFQSGRP-----VRRFVALAEQS 399
DB 290 EVDVCSHGR---ALMELSKVKEPPLWIKGNHNLIELYPEYIKHLKFV-TALEKS 341

RESULT 15
US-10-369-493-15631

Sequence 15631, Application US/10369493
Publication No. US20030233675A1
GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xianfeng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
FILE REFERENCE: 38-10(52052)B

;;
CURRENT APPLICATION NUMBER: US/10/369,493
CURRENT FILING DATE: 2003-02-28
PRIOR APPLICATION NUMBER: US 60/360,039
PRIOR FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO: 15631
LENGTH: 742
TYPE: PRT
ORGANISM: Xanthomonas campestris
US-10-369-493-15631

Query Match 4.8%; Score 104.5; DB 4; Length 742;
Best Local Similarity 21.0%; Pred. No. 0.88;
Matches 85; Conservative 50; Mismatches 161; Indels 109; Gaps 18;

QY 4 SASDGTHERSAPRLSPVLSGAMGPMHTGUYAQSMDVYLGQPDKLPARPTIALAAQ 63
DB 423 SASNGTHAAS-----FANNASVYVDTW-----SNTTTPQIB 454
QY 64 AFRDEIVLLGLKARRPVSNNRVERISQEVAAGLEFYGNRRMLEKPSGFFA-----QPP 118
DB 455 LFR-----ANGKIKATLLQNDVADPQHYPYAKYRQQRVEFGTLTAADGSTP 501
QY 119 LTBVAVRKXKDRRSFYRIFFDSGFTPHGEPSQRMSTYANNREYALLRHPRPWL 178
DB 502 L-----HYRLIKPAGFDP-----SKRY-----PVV 521
QY 179 VCVHGTGEMGRAPLDLAVFRAMKLDHDLGNIWVPLPMHGRGQLP-KGAVPGEVDLD 237
DB 522 VVYVGPPAAQTVVDAMPGRGDALFDQYLAQGYVVFSLD--NRGTPRRGRAFG--ALY 576
QY 238 DVHGTAAVMDIRRLSWIRSOE--ESLIGLNGLSLGGYIA-SLVASLEGLACAILGV 294
DB 577 GKQGYIE-VDDQGVAVMLKTQKRWVDARIIGVGNNGSGYMTMLAKHSEAYACGVAGA 635
QY 295 PVADLIELIGRRCGLRHKD-PRRHVYKMAP-IGRMISPLSTPLVPMGRFTYAGIADR 352
DB 636 PVTDM-GLYDTHYTERYMDLPAGNAGYREARIATHLDGLRAKL-----LIHGMADD 687
QY 353 LV---HPRQVTRLMEMHMKREIWMYRGHTGFFQSRPVRPFVQA 394
DB 688 NVLFTNSTALMSGLOQRGTPFELMTYPGAHGSLSGKTLALHRYXTA 732

Search completed: March 23, 2006, 06:28:19
Job time : 88.8513 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 23, 2006, 06:17:23 ; Search time 9.91096 Seconds
(without alignments)
1192.746 Million cell updates/sec

Title: US-10-617-038-29

Perfect score: 2195
Sequence: 1 MASSASDQTHRSAPRLSP.....AALEQSGILDAPRTGRDRA 413

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 169630 seqs, 28622889 residues

Total number of hits satisfying chosen parameters: 169630

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA_New.*
1: /cgn2_6/prodata/1/pubppaa/US08_NEW_PUB.pep.*
2: /cgn2_6/prodata/1/pubppaa/US06_NEW_PUB.pep.*
3: /cgn2_6/prodata/1/pubppaa/US07_NEW_PUB.pep.*
4: /cgn2_6/prodata/1/pubppaa/PCT_NEW_PUB.pep.*
5: /cgn2_6/prodata/1/pubppaa/US05_NEW_PUB.pep.*
6: /cgn2_6/prodata/1/pubppaa/US10_NEW_PUB.pep.*
7: /cgn2_6/prodata/1/pubppaa/US11_NEW_PUB.pep.*
8: /cgn2_6/prodata/1/pubppaa/US60_NEW_PUB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	98	4.5	668	6	US-10-454-437-118
2	96	4.4	2458	7	US-11-186-999-6
3	95	4.3	300	7	US-11-179-977-4
4	94.5	4.3	2455	7	US-11-186-999-4
5	93	4.2	702	7	US-11-096-568A-14646
6	93	4.2	1096	6	US-10-995-561-710
7	93	4.2	2458	7	US-11-186-999-13
8	92	4.2	402	7	US-11-096-568A-16978
9	91	4.1	2458	7	US-11-186-999-11
10	89.5	4.1	299	6	US-10-454-437-228
11	88	4.0	443	7	US-11-096-568A-9373
12	88	4.0	443	7	US-11-096-568A-9375
13	88	4.0	621	6	US-10-467-657-1554
14	87.5	4.0	253	7	US-11-096-568A-33023
15	87.5	4.0	383	7	US-11-096-568A-12774
16	87.5	4.0	406	7	US-11-096-568A-18614
17	87.5	4.0	453	7	US-11-096-568A-18613
18	87.5	4.0	455	7	US-11-096-568A-18612
19	87.5	4.0	475	7	US-11-087-099-2233
20	86.5	3.9	563	6	US-10-454-437-120
21	85.5	3.9	239	7	US-11-096-568A-33024
22	85	3.9	2455	7	US-11-087-099-5180
23	85	3.9	2455	7	US-11-186-999-14
24	85	3.9	2455	7	US-11-186-999-16
25	84.5	3.8	383	7	US-11-096-568A-5747

26	84.5	3.8	2256	7	US-11-144-368-4	Sequence 4, Appli
27	84	3.8	298	7	US-11-179-977-6	Sequence 6, Appli
28	84	3.8	2348	6	US-10-450-224A-2	Sequence 2, Appli
29	83.5	3.8	264	7	US-11-096-568A-21268	Sequence 21268, A
30	83.5	3.8	297	7	US-11-096-568A-21267	Sequence 21267, A
31	83.5	3.8	674	7	US-11-167-048-1	Sequence 1, Appli
32	83	3.8	213	6	US-10-467-657-3394	Sequence 3394, Ap
33	82.5	3.8	366	6	US-10-329-258-27	Sequence 27, Appli
34	82.5	3.8	366	7	US-11-000-463-410	Sequence 410, App
35	82.5	3.8	366	7	US-11-000-463-882	Sequence 882, App
36	82.5	3.8	405	7	US-11-087-099-1566	Sequence 1566, Ap
37	82.5	3.8	407	7	US-11-087-099-5591	Sequence 5591, Ap
38	82.5	3.8	1057	6	US-10-330-773-253	Sequence 253, App
39	82	3.7	334	7	US-11-096-568A-10349	Sequence 10349, A
40	82	3.7	371	7	US-11-096-568A-10348	Sequence 10348, A
41	82	3.7	3073	7	US-11-143-980-50	Sequence 50, Appli
42	82	3.7	5712	7	US-11-143-980-47	Sequence 47, Appli
43	81.5	3.7	370	7	US-11-096-568A-28963	Sequence 28963, A
44	81.5	3.7	377	7	US-11-096-568A-28962	Sequence 28962, A
45	81.5	3.7	392	7	US-11-096-568A-29673	Sequence 29673, A

ALIGNMENTS

RESULT 1
US-10-454-437-118
; Sequence 118, Application US/10454437
; Publication No. US20050277115A1
; GENERAL INFORMATION:
; APPLICANT: Pompeius, Markus
; APPLICANT: Kroege, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Habberhuer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
; FILE REFERENCE: INVOLVED IN HOMEOSTASIS AND ADAPTATION
; CURRENT APPLICATION NUMBER: US/10/454,437
; PRIOR FILING DATE: 2003-06-13
; PRIOR APPLICATION NUMBER: US 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: DE 19931636.8
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19932125.6
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932126.4
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932127.2
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932128.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932129.9
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: DE 19932226.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932920.6
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932922.2
; PRIOR FILING DATE: 1999-07-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 442
; SEQ ID NO 118
; LENGTH: 668
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
; US-10-454-437-118
Query Match 4.5%; Score 98; DB 6; Length 668;
Beef Local Similarity 20.8%; Pred. No. 0.59;
Matches 91; Conservative 52; Mismatches 157; Indels 138; Gaps 25;
QY 41 RDYIGQDPKLPINRPTIALAAQAF-----RDEIVLVGLKRRPVSNHR- 84

```

Db      289 KEFLGCDRGQVFFPPTTESTSLQGLATTKNPLVLTILNNVSTIEIVPL--NDPTTEHH 346
Qy      85 --VFERSQEVAAGLEFYGNRRMLEKPSGFPAQPPLTev---AVRKVDRRRSPYRLF 138
Db      347 IDLPBHVAHVAVATSPBDGDEIWAQAA--FTEAPLTLLRAELPGALBAVKKAPLQF---- 400
Qy      139 FDSGFTPHGEESQSMLSYTNAREYALLR---HPEPRLVVCVCHGTMGAPLDAV 195
Db      401 -----ENAGGQETRHQWAT--SADGTXI PYFITGAPEEPDPNTLVHAYG----- 441
Qy      196 FRWKLHDELGNITWVPVPMHGR--RG--QGLPKGAVF-----PGEDVDVHGTA--- 243
Db      442 -----GFEXSLTPSHSPYTGIMLEGGYFVEANLRGGSGPGEWMSQATKLN 489
Qy      244 -QAVV--DIRRLTSMI--RSQEESLIGLNGLSGTYAS--LVASLEGALCAITLGPVAD 298
Db      490 RKKWEDHRAVLAIDLVERGYATPEQJLAITGGSGNGLTSGALTQYFEARCAAVVQPLAD 549
Qy      299 LIELGRHGLAKKDPRRRTVK-----MAE-----PIGRMI--SPLSLTPVP--M 340
Db      550 ML-----RHYTWSAGASMAEYGNPDDEPBEBAVIEQISPVQAVVGEKRIY 595
Qy      341 PGRPLVIAGIADRLVHPREQVTRIM-----EHMGKPEIWMYPGCHTG-----F 382
Db      596 PPLATVTTSTRDVRVPAH--ARLFAQALIDAGQAVDYENTE-----GGHAGADNKGTA 648
Qy      383 FQSRVRRFVQAALQSG 400
Db      649 FVESLITWIEKTLDOQG 666

```

```

RESULT 2
US-11-186-999-6
; Sequence 6, Application US/11186999
; Publication No. US20060019364A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: ACETYL COA CARBOXYLASE 2 SEQUENCES AND METHODS
; FILE REFERENCE: BMS Dossier Number 10245 NP
; CURRENT APPLICATION NUMBER: US/11/186,999
; CURRENT FILING DATE: 2005-07-21
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6
; LENGTH: 2458
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-186-999-6

Query Match          4.4%; Score 96; DB 7; Length 2458;
Best Local Similarity 20.6%; Pred. No. 5,2;
Matches 93; Conservative 46; Mismatches 131; Indels 182; Gaps 23

QY      37 AQSMDYL--GQDPKLPRIARPTLAAQAFDEIVLGL-----KRRPVSHRVER 88
Db      198 ALSLEAVLTGAEATRVFTMRSPMS-----GLHLVKRGREHKKLDLRDF-- 242
      198 ALSLEAVLTGAEATRVFTMRSPMS-----GLHLVKRGREHKKLDLRDF-- 242

QY      89 ISQEVAAQLER---YGNRRMLEK---PSGFPAQPEPLTEVAVRKYQRRRSFYIF-- 138
Db      243 ---TVASPAEFTVRYGSGDRVIEKVIANNNGI-----DAVKCRSIRRWAYEMFRNE 290
      243 ---TVASPAEFTVRYGSGDRVIEKVIANNNGI-----DAVKCRSIRRWAYEMFRNE 290

QY      139 -----FDSGFTPHGPEPSQRMVSLYTANNREYA-----DLL 169
Db      291 RAIFPVVWVTEDDLKANAETIKMADHYFVPGGP-----NNNNYANVELLVDIATK 340
      291 RAIFPVVWVTEDDLKANAETIKMADHYFVPGGP-----NNNNYANVELLVDIATK 340

QY      170 RHPEPRPV-----LYVCVHSTEMGRAPLDIAVFRAMKLDDELGINIV----- 210
Db      341 RIFPRAVWAGHGASENPKLPBLLCKNGVAFLGPSE-----AMMALGDKRIASVVAQTLL 396
      341 RIFPRAVWAGHGASENPKLPBLLCKNGVAFLGPSE-----AMMALGDKRIASVVAQTLL 396

QY      211 MPVLPNMGPR-----RGQGLPKNG-AVFQGEIVLD-----DVH----- 240
Db      397 VFTLPWSSGGLTVENTEDDLQGGKRISVPEDEVYDKGVCVADDEGLEAERIGFPLMIKAS 456
      397 VFTLPWSSGGLTVENTEDDLQGGKRISVPEDEVYDKGVCVADDEGLEAERIGFPLMIKAS 456

```

[illegible]

```

RESULT 3
US-11-179-977-4
: Sequence 4, Application US/11179977
: Publication No. US20050245789A1
: GENERAL INFORMATION:
: APPLICANT: Genencor International, Inc.
: TITLE OF INVENTION: Alpha/Beta Hydrolase-Fold Enzymes
: FILE REFERENCE: GC511-PCR
: CURRENT APPLICATION NUMBER: US/11/179,977
: CURRENT FILING DATE: 2005-07-12
: NUMBER OF SEQ ID NOS: 21
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 4
: LENGTH: 300
: TYPE: PR1
: ORGANISM: Bacillus
PS-11-179-977-4

```

```

Query March 4.3%, Score 95; DB 7; Length 300;
Best Local Similarity 23.1%, Pred. No. 0.38;
Matches 56; Conservative 41; Mismatches 109; Indels 36; Gaps 11

QY      171 HPEPRMLVCVHGTEMGKRAPIDLAIVFRAWKHDELGLNIYMPVLPMHGPRGQLEKGAIV 230
           ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB      78 HDFTNTITIC-HGVTMN---VLNLSLKVMHFLFDLGNVNLIIYDRHRHQSG-KGTTSYGF 131

QY      231 PGEIVLDDVHGTAAVMDIRLLSWISSQ-EESDLGNGISLGVIASLVA--LEBGL 287
           :|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB      132 YEKD-----DLNKVSLKXKTNHKGIGIHGSGMGAVTLLLYAGHCSGGA 178

QY      288 ACAIIGVPVADLIELGRHCGLRKDRPRHTVKAEPi-----GRMISPLSTPLVP 339
DB      179 DFYIADCFPCDFDQALVRLAEFRRLSPPLPIADFLKLKRGVRAVEGPLAVIDKIE 238

QY      340 MPGEFIYAGIADRLVHPREQVTRIMHEWKGPEIWWP--GGHT-GFQSR-PVRRFVQA 395
DB      239 KPVLFHISKDDYI--EVSSTERLYEKKRGKPKALYIAENGCHAMSYTKNRHTYRKTQEF 296

QY      396 LE 397
           :|
DB      297 LD 298

RESULT 4
US-11-186-999-4
; Sequence 4, Application US/111866999
; Publication No. US20060019364A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: ACETYL COA CARBOXYLASE 2 SEQUENCES AND METHODS
; FILE REFERENCE: BMS Docket Number 10245 NP
; CURRENT APPLICATION NUMBER: US/11/186,999
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4
; LENGTH: 2455
; TYPE: PRT
; ORGANISM: Homo sapiens

```

US-11-186-999-4

Query Match

4.3%; Score 94.5; DB 7; Length 2455;

Best Local Similarity 20.5%; Pred. No. 7.1;

Matches 92; Conservative 47; Mismatches 131; Indels 179; Gaps 23;

```
QY 37 AQSMDYL--GOQPDKLPARPTIALAAQAFDEIVLGL-----KARRPVSNHRVPER 88
DB 198 ALSLAEVLTGTAETRVFTMRPSMS-----GLHLVKRGREHKLDLRDF-- 242
QY 89 ISOEVAAGLEF---YGNRRMLEK---PSGFFAQPPLTEVAVRKVKDRRRSFYRIF--- 138
DB 243 ---TVASPAEFVTRPGGDVIEKVLANNGI-----AAVKCRSIRRAWAYEMFRNE 290
QY 139 -----FDSGFTPHGPEPGSQRWLSYANNREYA-----LLL 169
DB 291 RAIRFVVMVTPEDLKANAEYIKMADHYFVPGGP-----NNNNYANVELLVDIAK 340
QY 170 RHPEPRPW-----LVCVHGTEMGRAPLDLAVFRAMKLDHGLGNIYV-----MPV 213
DB 341 RLPVQAGWGHASBNPKLPBILCKNGVAFLGPSE-----AMWALGDKIASVVAQTLQVPT 396
QY 214 LPMHGP-----RGQGLPKG-AVFPGEDYLD-----DVH----- 240
DB 397 LPMSSGSLTVENTEDDLOQGKRI SVPEDDYDKGCVDVDEGLEAERIGPPLMITKASGG 456
QY 241 -----GTAQAVNDIRLLSMIRSOEBSLIGLNGSLGGYIASLVASLEEGACAILGVP 295
DB 457 GKGIRKASADPFLFRQVQSEIRGSPILMKL-----AQHARHLEVOQLADQ 506
QY 236 VADLIELGRHCGLRHKDRRHTVKAEPICGMISLSTPLVPMG-----RTIYVG 348
DB 507 YGNAVSFLGRDCSIQ---RRHQ-KIVEBAPATIAFLAIFEMEQCAIRLAKTVGVVSAG 561
QY 349 IADRL-----VHREOVTRLMWH 366
DB 562 TVEYLYSQDGSFPHFELNPRLOY---EH 586
```

RESULT 5

US-11-096-568A-14646

Sequence 14646, Application US/11096568A

Publication No. US20060048240A1

GENERAL INFORMATION:

APPLICANT: Alexandrov, Nikolai et al.

TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides

FILE REFERENCE: 2750-1592PUS2

CURRENT APPLICATION NUMBER: US/11/096,568A

CURRENT FILING DATE: 2005-04-01

NUMBER OF SEQ ID NOS: 34471

SEQ ID NO 14646

LENGTH: 702

TYPE: PRT

ORGANISM: Zea mays subsp. mays

FEATURE:

NAME/KEY: misc.feature

LOCATION: (1) --(702)

OTHER INFORMATION: Ceres Seq. ID no. 11049476

US-11-096-568A-14646

Query Match

4.2%; Score 93; DB 7; Length 702;

Best Local Similarity 23.6%; Pred. No. 1.8;

Matches 109; Conservative 35; Mismatches 171; Indels 146; Gaps 23;

```
QY 5 ASDGTHESAFPLSPVLSGAMGPFMHTGLVYAQSRDYLGOQPDKLPARPTIALAAQ 64
DB 212 AADTGRRRRRLRLP---LHGDRAP-----APGARD---RPPRLPGQH---WFAADA 253
QY 65 PFDEIVLGLKARR-----PVSNHRVFERISOEVAAGLEFYGNRRRLKXSPGFAQ---P 116
DB 254 LBDRAPRVRLARRRVRPRGRLGHRGRQ-----RPLLRQPRPAHVRGRPP 300
```

QY

117 PRLTEVAVRKVDRR-----RSFYRIFFDSGFTPHGPEPGSQRWLSYANNREYALLRHP 172

DB

301 LRPFRGRRPRDRRLRLRRARLHDDAPAGPGRKAPR-----AQLRRNQ 349

QY

173 EPRPLVVCY---HSTEMGRAPLDLAVFRAMKLDHGLGNIYMPVLMHGP--RGQGLPKG 227

DB

350 EAVFOVLLQARRHQVRRRGDPAGAA-----HDD-----PRLRHHRELRCARPPG 395

QY

228 AVFPGEDVLDVHGTAQAVNDIRLLSMIRSOEBSLI-----GLNGLSGGYIASL 279

DB

396 GV-----QAPGDAARRVARGAGGGEVUAVARAPQARRGRGVGGAGLIL 441

QY

280 VASLEB-----GLACAILGVVADLIELGHHGGLR--HKDP--RRH----- 317

DB

442 LILPVEYVGRGDRGGGDRLLHDPRLHLQVRRAPEGRADRDPPGRAHGFVAHPRPA 501

QY

318 TVKMAEPICGMISPLTP---LVPMEGRFYAGIADRL---VHREOVTRLMWHGKP 370

DB

502 AVAAGEPGGGHEPFPAGPREVYRPRGGGAVAGVGLRGSPVIGRAHQVAV----- 554

QY

371 EIVWYPGGHTGFPQGRPVYRQVQALQSGLLDAPRTQDR 411

DB

555 -----RGGPYRR--RALLRSHGPRGPRARRGR 580

RESULT 6

US-10-995-561-710

Sequence 710, Application US/10995561

Publication No. US20050272054A1

GENERAL INFORMATION:

APPLICANT: CARGILL, Michele et al.

TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH

TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF

FILE REFERENCE: CLO01559

CURRENT APPLICATION NUMBER: US/10/995,561

CURRENT FILING DATE: 2004-11-24

NUMBER OF SEQ ID NOS: 85702

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 710

LENGTH: 1096

TYPE: PRT

ORGANISM: Homo sapiens

US-10-995-561-710

Query Match

4.2%; Score 93; DB 6; Length 1096;

Best Local Similarity 20.4%; Pred. No. 3.3;

Matches 92; Conservative 47; Mismatches 131; Indels 182; Gaps 23;

```
QY 37 AQSMDYL--GOQPDKLPARPTIALAAQAFDEIVLGL-----KARRPVSNHRVPER 88
DB 198 ALSLAEVLTGTAETRVFTMRPSMS-----GLHLVKRGREHKLDLRDF-- 242
QY 89 ISOEVAAGLEF---YGNRRMLEK---PSGFFAQPPLTEVAVRKVKDRRRSFYRIF--- 138
DB 243 ---TVASPAEFVTRPGGDVIEKVLANNGI-----AAVKCRSIRRAWAYEMFRNE 290
QY 139 -----FDSGFTPHGPEPGSQRWLSYANNREYA-----LLL 169
DB 291 RAIRFVVMVTPEDLKANAEYIKMADHYFVPGGP-----NNNNYANVELLVDIAK 340
QY 170 RHPEPRPW-----LVCVHGTEMGRAPLDLAVFRAMKLDHGLGNIYV----- 210
DB 341 RLPVQAGWGHASBNPKLPBILCKNGVAFLGPSE-----AMWALGDKIASVVAQTLQ 396
QY 397 VFTLPWSSGSLTVENTEDDLOQGKRI SVPEDDYDKGCVDVDEGLEAERIGPPLMITKAS 456
DB 241 MEVLPMHGP-----RGQGLPKG-AVFPGEDYLD-----DVH----- 240
QY 241 -----GTAQAVNDIRLLSMIRSOEBSLIGLNGSLGGYIASLVASLEEGACAIL 292
DB 457 BGGGKGIRKASADPFLFRQVQSEIRGSPILMKL-----AQHARHLEVOQL 506
```

Qy 253 GVPVADLELLGRHGLRHKDPRRHVYVKAEPYIGRNISPLSLNPLVPMW-----RF1 345
Db 507 ADQYGNVSVLFRDGSIQ-----RRHQ-KLYEENPATIAPLAIIEFMGQCRLAKTVGY 551
Qy 346 YAGIADRL-----VHPPEQVYTRIMEH 366
Db 562 SAGTVEYLYSGDGSFHFLELNPRLQV-----EH 589

RESULT 7
US-11-18

US-11-186-999-13
; Sequence 13, Application US/11186999

```

: GENERAL INFORMATION:
: APPLICANT: Bristol-Myers Squibb Company
: TITLE OF INVENTION: ACETYL COA CARBOXYLASE 2 SEQUENCES AND METHODS
: FILE REFERENCE: BMS Docket Number 10245 NP
: CURRENT APPLICATION NUMBER: US/11/186,999
: CURRENT FILING DATE: 2005-07-21
: NUMBER OF SEQ ID NOS: 34
: SOFTWARE: PatentIn version 3.2
: SEQ ID NO 13
: LENGTH: 2458
: TYPE: PR1
: ORGANISM: Homo sapiens
: US-11-186-999-13

```

Query Match	4.2%	Score 93	DB 7	Length 2458
Best Local Similarity	20.4%	Pred. No. 9.8		
Matches 92; Conservative	47	Mismatches 131	Indels 182	Gaps 23

```

QY 37 AQSMDYU---GGQBDLPTAAPTIALAQAQFREIYLLG-----KARPVSNHVFER 86
Db 198 ALSLFAVLLTGAETRVPTMRPSMS-----GLHVKRGREHKGLDHRDF-- 242
QY 89 ISOEVAAGLEF---YGNRRMLEK---PSGFAOPPLTEVAARKVRRRSFYRIF-- 138
Db 243 ---TVASPAEFVTRGGDRVIEKVLIANNGI-----AAVKCMRSIRMAVEMFRNE 290
QY 139 -----FDSGFTHPFGPSGRMLSTYANNREYA-----LLL 169
Db 291 RAIRFVWVTPEDLKANAEYIKMADHYVVPGRP-----NNNNVANVELVDIAK 340
QY 170 RHPEBRPW-----LVCHGTGEMGAPLDLAVFRAMKJHDELGLNIY----- 210
Db 341 RLPQVAWAGMGHASENPYLPBELLCKNGVAFLGPRSE---AMMALGDKIASIVVAQTLQ 396
QY 211 MPVLHMGP-----RGQGLPKG-AVFPGEVDLD-----DVH----- 240
Db 397 VPTLFMSGSGLTVEWTEDDLQOGKRISEBEDYVDKCVKDVBREGLAEERIGEPMLIKAS 456
QY 241 -----GTNAQVMDIRRLBSMIRQEBESLIGLNGLSIGYIATLVASLEBGLACATL 292
Db 457 EGGGSGKIRKASAEFPILPROVOSEIRGSPFLMKL-----KOHARHLEVQIL 506
QY 293 GVPVADLLIELGRHCGGLRHKDPRRRHVKKAAEPIGRMISPLSLTPLYMPBP-----RPI 345
Db 507 ADQYENAVASLFRDCSIQ---RRHQ-KIVEEAPATITAPLAIEFMEQCARILAKTVGVY 561
QY 346 YAGIADRL-----VHREQYUURLMEH 366
Db 562 SAGTVEVLYSODGSFHFLEINRLQV-----EH 589

```

RESULT 8
US-11-09

US-11-096-568A-16978
Sequence 16978. Application US/11096568A

PUBLICATION NO.: US20060046240A1
 GENERAL INFORMATION:
 APPLICANT: Alexandrov, Nikolai et al.
 TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
 TITLE OF INVENTION: Theory
 FILE REFERENCE: 2750-1592PUS2

```

: CURRENT APPLICATION NUMBER: US/11/096,568A
: CURRENT FILING DATE: 2005-04-01
: NUMBER OF SEQ ID NOS: 34471
: SEQ ID NO 16978
: LENGTH: 402
: TYPE: prt
: ORGANISM: Zea mays subsp. mays
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (1)..(402)
: OTHER INFORMATION: Ceres Seq. ID no. 123556322
US-11-096-568A-16978

```

Query Match	4.2%;	Score 92;	DB 7;	Length 402;
Best Local Similarity	22.3%;	Pred. NO. 1.1;		
Matches 81; Conservative	40;	Mismatches 134;	Indels 108;	Gaps 17

```

Qy      39  SMDYILGGQPDKLPFLARPTIALAQA-----FRDIYVLGKKAR----PVSNRVVF  86
Db      21  SSRDTPAOPORRCVAFPVROMAASATADAGAGLRSEFLOVLLSRRDDLOVPLS-----  75
Qy      87  ERISOEVAAGLEFYGNRWLKERPSGFPAOP--PLTEAVARKYCDRRRSFYRIFFDSCF  143
Db      76  --VEGSSPYKXDPMYQN-----PPTIMEANPMSPCSKOVENLRKLVBEBNYLL-----  122
Qy      144  TPHPGSPGSQRMWLSYTNANREYALLLRHPD--RPLVLCVHGCT---EKGRAPDLAV  195
Db      123  -TELGEQG-----RVPVLLLLKLDDPVPRRKRPAIVFLHSSYCKCKEMLLPRLLEYA  170
Qy      196  FRW-----KLHDELGLNTVMPLRPHNGRGCGLPKGAVFPEEDVLDDVHGAQAQVMD  248
Db      171  SKSYCVVAIDSRVHGGRASNETTYIEALKSAMRNG-----DTMPFLFDI---VMD  217
Qy      249  IRLLSMIRSOE--EESLIGLNGLSGLGYIASLVLSLEE--GLACATILGV-----  294
Db      218  LVKLGSYLGEREVDVPSRGTGSELSGGMHMFAPFVDTKRSVVVPLIIGVQPRMADNN  277
Qy      295  -----PVADLIETLGRHCGLRHKDPRRHVTYKQAEPIGRMI-----SPLSLTPLV  358
Db      278  KMLARVNSIRLPFEAKGIDLGKSELD-----TEVEKWEKXIAPGLDSQFDAPAPSLPLA  332
Qy      339  PMP 341
Db      333  PRP 335

```

RESULT 9
US-11-18

US-11-186-999-11
Sequence 11, Application US/11186999

```

: GENERAL INFORMATION:
: APPLICANT: Bristol-Myers Squibb Company
: TITLE OF INVENTION: ACETYL COA CARBOXYLASE 2 SEQUENCES AND METHODS
: FILE REFERENCE: BMS Docket Number 10245 NP
: CURRENT APPLICATION NUMBER: US/11/186,999
: CURRENT FILING DATE: 2005-07-21
: NUMBER OF SEQ ID NOS.: 34
: SOFTWARE: PatentIn version 3.2
: SEQ ID NO 11
:
: LENGTH: 2458
:
: TYPE: PRT
:
: ORGANISM: Homo sapiens
:
: FEATURE:
:
: NAME/KEY: MISC FEATURE
: LOCATION: (1)..(2458)
:
: OTHER INFORMATION: Xaa is any amino acid
US-11-186-999-11

```

Query Match	4.1%	Score 91;	DB 7;	Length 2458;
Best Local Similarly	20.9%;	Pred. No. 15;		
Matches 96; Conservative	55;	Nonmatches 165;	Indels 144;	Gaps 23;

QY 1 MASSASDGTHERS-----AFRLSPVLG-----AMGPFMHTGLVVAQSWRDYLGQP 48


```

Db      180 VASSSSSESTRKGRASLGLSLSEAYLITGAEATRVPTMRPSM--SGHLVLR-----GREH 233
Qy      49 DKLPILRP--TIALAAQAFRDEIVLGLKARRPVSNHVFPERI---SQEVAAGLEFYGNRR 104
Db      234 KKLIDLRDPTVASPAEF-----VTRXG-----GRRVIEKVI IANNGLAAVCMGSIIR 281
Qy      105 WKEKPSGFFAQPPLTEVAARVKYKDRRRSFYRIFFDSGFTPHGEGSQWMLSYTANNE 164
Db      282 MAYE---MFEENRAIRFVVVWVTPEDLKANAKEYIKMADHYVPVGGP-----NNNN 328
Qy      165 YA-----LLLRHPRPRVLVCVHGTENGRAP-----LDLAVPRAMKHLDEL 207
Db      329 YANVELIYDIARKIIPYAXAMKXXXXHASSEMPKLPBELLCKNGVAFLGPPSEAMMALGDKIAS 388
Qy      208 NIY-----MPVLPMPHP-----RGQGLPKG--AVFPGEDVLD-----DYH----- 240
Db      389 TVVAQVLQVPTLPWSSSGLTVENTEDDLOQSKRISPEDEYDGCYKVDDEGLBAERIG 448
Qy      241 -----GTAQAVMDIRLLSMIRSOEESLIGLNGSLGQYIASLVASLE 284
Db      449 PKLMIKASGGGKGIRKAKESADPPLIFRQVQSEIRGSPILMKL-----AQHA 498
Qy      285 EGLACAILGVPAVDLELGRHCGLRKDPRTTVMAEPFGMISPLTFLVMPG-- 342
Db      499 RHLVQVILADQYGNNAVSLFGRDCSIQ-----RRHQ--KIVEAPATIAPLAIFEFMEQCAIR 553
Qy      343 -----RFIYAGIADRL-----VHPREOVTLMEH 366
Db      554 LAKTVGYVSGKXVEYLYSQDGSFHPLELNRLOY-----EH 589

```

RESULT 10

```

US-10-454-437-228
; Sequence 228, Application US/10454437
; Publication No. US20050277115A1
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroeger, Burkhard
; APPLICANT: Schroeder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Habernauer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
; FILE REFERENCE: BGI-128CPCN
; CURRENT APPLICATION NUMBER: US/10/454,437
; CURRENT FILING DATE: 2003-06-13
; PRIOR APPLICATION NUMBER: US 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: DE 19931636.8
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19932125.6
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932126.4
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932127.2
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932128.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932129.9
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: DE 19932226.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932920.6
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932922.2
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 442
; SEQ ID NO 228
; LENGTH: 299
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum

```

US-10-454-437-228

Query Match

4.1%; Score 89.5; DB 6; Length 299;

Best Local Similarity 24.9%; Pred. No. 1.2;

Matches 63; Conservative 25; Mismatches 114; Indels 51; Gaps 10;

```

Qy      176 PMLVCHGTENGRAPLDLAVPRAMKHLDELGINV-----MPVLPMPHGGGLPKGAVF 230
Db      63 PCLVYFHGGGMSGGTLNMDATVHSL--VVGPIIAISVDYLAAPH----- 107
Qy      231 PGEDVLDVHGTAAQVMDIRLLSMIRSOEESLIGLNGSLGQYIASLV--SLEEG--- 286
Db      108 PEPALDDAFVAVSAVLDSVGLS-----IDSRVAIGDSAGNIAAVTAQQLERANG 162
Qy      287 ---LACAILGVPAVDLE-----LGRHCGLRKDPRTTVMAEPDGR---MISPL 332
Db      163 STPVLAAQVLIFPVTDVSTSTPSYLPFGKOCYGLTKDAMERYIEQYADGHDRTDPLSP 222
Qy      333 SLTPVMPMGRTIYAGIADRLVHP--REOVTRLMEHMKREIYVYRGHTGFF-----Q 384
Db      223 LASDLSDEPPTIYVGECDVLAHEVRAYGOALLLEAGNSVTMTPEFKQIHAFINLGSISS 282
Qy      385 SRPVRRFVQALE 397
Db      283 ARPARRLRAELE 295

```

RESULT 11

```

US-11-096-568A-9373
; Sequence 9373, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592FUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 9373
; LENGTH: 443
; TYPE: PRT
; ORGANISM: Triticum aestivum
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 11--(443)
; OTHER INFORMATION: Ceres Seq. ID no. 14302452
; US-11-096-568A-9373

```

Query Match 4.0%; Score 88; DB 7; Length 443;

Best Local Similarity 22.4%; Pred. No. 2.9;

Matches 83; Conservative 37; Mismatches 115; Indels 136; Gaps 21;

```

Qy      44 LQQQDPKLPAPPTIALAAQAFRDEIVLGLK--ARRPVSNHVFERISQEVAA----- 95
Db      11 VQGLGRIRLPFGAS-----ARKAFHCEIRLRGFPVQLAPVPLENSSEPHLDPHANAAVFSLD 65
Qy      96 --GLEFYGNRRWLEKPSGFFAQPPLTEVAV-----RKYKDRRRSFYRIFFDSGFT 144
Db      66 EGGLEALPTSRWFRPTE-----PYLETAIVSVSRDGGGCHGCLRRLVGAVRVDVGE 119
Qy      145 PHRPGSGQWMLSYTANNREYALLRHPERPMVLVCVHGTENGRAP-----LDLAVPR 197
Db      120 WRDQKP-----VLRH-----SW-----TGIGRAELHVRVRVEADPRY 153
Qy      198 AKKLHDELGLNIVMPVLPMPHPRGGGLPKGAVFPGBEDVLDVHGTAAQVMDIRLLSMIR 257
Db      154 VRFPEDEVALN--PQVQLHGGASQ-----PIF-----SCKFIQVDVR-----A 190
Qy      258 SOEBSLIGLNGSLGQYIASLVASLEGLACAILGVPAVDLELGRHCGLRKDPRTTVMA 317
Db      191 SQPDH-----LGKTVSSSSGSGEE-----KETENAG-----RRERKCM 224

```

Qy 318 TVKMAEPIGRMSPSL-TPLVPM-----PGRFTYAGIADLVHPREQV---RL 363
Db 225 KVAIHDLSSAAVAAMFMAFPVPAIGSDTVARSNPQAMLI-----RADTTSSSS 275
Qy 364 WEHQGKPELIV 374
Db 276 WQPMGRLE-AW 285

```

RESULT 12
US-11-096-568A-9375
; Sequence 9375; Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 9375
; LENGTH: 443
; TYPE: PRT
; ORGANISM: Triticum aestivum
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(443)
; OTHER INFORMATION: Ceres Seq. ID no. 1661697
US-11-096-568A-9375

```

Query Match	4.0%;	Score 88;	DB 7;	Length 443;
Best Local Similarity	22.4%;	Pred. No. 2.9;		
Matches	83;	Conservative	37;	Mismatches 115;
				Indels 136;
				Gaps 21;

[illegible]

RESULT 13
US-10-467-657-1554
Sequence 1554, Application US/10467657
Publication NO. US20050260581A1
GENERAL INFORMATION:
APPLICANT: CHIRON SPA
APPLICANT: FONTANA Maria Rita
APPLICANT: PIZZA Mariagrazia
APPLICANT: MASIGNANI Vega

```

:
: APPLICANT: MONACI Elisabetha
: TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
: FILE REFERENCE:
:
: CURRENT APPLICATION NUMBER: US/10/467,657
: CURRENT FILING DATE: 2003-08-11
: PRIOR APPLICATION NUMBER: GB-0103424.8
: PRIOR FILING DATE: 2001-02-12
: NUMBER OF SEQ ID NOS: 9218
: SOFTWARE: SeqMan9, version 1.04
: SEQ ID NO 1554
:
: LENGTH: 621
:
: TYPE: FRT
:
: ORGANISM: Neisseria gonorrhoeae
:
: US-10-467-657-1554

```

Query Match	4.0%;	Score 88;	DB 6;	Length 621;
Best Local Similarity	25.2%;	Pred. NO. 4.5;		
Matches	53;	Conservative	29;	Mismatches 74;
				Indels 54;
				Gaps 13;

```

QY      211  MPVLPMH--GPRQG--GLPKCAVPGSDVLDYDHNGTA-----QAVMDIIR-----LSW 255
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      404  IPELPIHYLOSVCQKYLMEENAF---VLANTIRGGGEFGFRMNDAAQGISKHKSVDDLLAV 459
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY      256  IRSQEEESL-----IGLNGLSLGGYIASLVASLE--EGLACAILGVPVADIIEI-----LGR 305
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      460  VRDLSRSGMSSPHNIGLGGSGNGGLITMAFVREPGQSIGALVCEVPLTDMIRPLLSAGS 519
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY      306  HCGLRHKDPRRRTVKKAEPTGMMISPLS-----LTPVPMGFRFYAGIADRLVNP----- 356
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      520  SMTDEVGNFOKY-----EACKRRLEOLSPRYNHLSDGIDYPRPALITLSDDRVHPALAK 574
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY      357  -----RGQVTRLMHHNGKEIIVWPDGGHTG 381
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      575  FYAKLRETSPQSWLY--SPD-----GGGHTG 598
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

```

```

RESULT 14
US-11-096-568A-33023
; Sequence 33023, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Thereby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 33023
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)..(253)
; OTHER INFORMATION: Ceres Seq. ID no. 13601781
US-11-096-568A-33023

```

Query Match	4.0%;	Score 87.5;	DB 7;	Length 253;
Best Local Similarity	21.8%;	Pred. No. 1.5;		
Matches	43;	Conservative	28;	Mismatches 73;
				Indels 53;
				Gaps 8;

Qy 117 PPTLEAAVAKVDORRRSFRI-----FPSS---GFPHPGEGSGSRMISLYANREYA 166
 Db 5 PPSRSLSVSMAD--SAFRKIQIGRDPTTFDAYVVGKODAPGIVLIGEMWGVPPEINHA 62
 Qy 167 LLLRHREPRRLVCVCHGTEMGARLDLAVFPAKMLHBLGINTMPLPHNPGRGGLPK 226
 Db 63 IKISQLEP-GFKALIPDLRYGKVGSDTA-----EAQHLM 96
 Qy 227 GAVPGEDVLDDVHGTAQVMDIRLLSLWIRSQEBSLIGLNGSLCG--YIASLVASLE 284
 Db 97 GIDWPG-----AIKIDIRASVVMKLSNDSKK-VGVTGCMCGAGALAISSVLVE 143

QY	285	EGLACAILGV	PVADLIE	301
		:	:	
Db	144	VDVVGGFYGT	PPSSELD	160

```

RESULT 15
US-11-096-568A-12774
: Sequence 12774, Application US/11096568A
: Publication NO. US20060048240A1
: GENERAL INFORMATION:
: APPLICANT: Alexandrov, Nikolai et al.
: TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
: TITLE OF INVENTION: Theory
: FILE REFERENCE: 2750-1592PUS2
: CURRENT APPLICATION NUMBER: US/11/096, 568A
: CURRENT FILING DATE: 2005-04-01
: NUMBER OF SEQ ID NOS: 34471
: SEQ ID NO 12774
: LENGTH: 383
: TYPE: PRT
: ORGANISM: Trifolium aestivum
: FEATURE:
: NAME/KEY: misc feature
: LOCATION: (1)..(383)
: OTHER INFORMATION: Ceres Seq. ID no. 14303558
: US-11-096-568A-12774

```

Query Match	4.0%;	Score 87.5;	DB 7;	Length 383;
Best Local Similarity	23.2%;	Pred. No. 2.6;		
Matches	78;	Mismatches	137;	Indels 75; Gaps 19;

```

Cy      89  ISEVVAAGLEFYENRRMLEKSGFQOPPLTEVAVRKVKRRSFRVIFPDSGFTBPG 148
      :: ||| ::
Db      4  VASTVAARAFRFP-----PSPEYGVPEPPSSAADA-----DAVVELSG 44

Cy      149  EPG-----SQRMSTYANNREYALLRHPEPRMVLVCNG--TEMGRAIDLAVPRAKLN 202
      :: ||| ::
Db      45  VPRGGVEARRLEPTKRSGSEVAMAYAKQAPARLTLLYSHGNADJGOM-YELFV----ELS 99

Cy      203  DEKGLNIWMPVLPMHGPQGLPKGAVFGEVDLVDVGTAAVMDIRLLSWIRSOEE 262
      :: ||| ::
Db      100  SHLWNLN-----GYDVSIGYGQSSGKRSE---QNTADIEAV--RCLIEYGASEE 166

Cy      263  SLGLNGLSIGYITASLVASLEBGLACAILGVPAVDLIELGRHCGLRHNDPRHYYKMA 322
      :: ||| ::
Db      147  NII-LVGOSGSGSPETLDLASRLPHLRVAVLAHSPIS-----GLRWMPVGHYTF- 195

Cy      323  EPIGRMISPLSLTPVPMGFRFYAGIADRLV---HPREQVTRLME-HMGKPIAVIYPGG 378
      :: ||| ::
Db      196  ----DIYKIDKVALKCFVLVITH-GTADVDVDCSHGR-----ALMELSKYIEPLWYKGG 246

Cy      379  HTEGFGOSRP-----VRRFQALAEQSGLL--DAPRT 407
      :: ||| ::
Db      247  NHCNLELVPEYIYHMLKKFI-TAIEKSPPLKDDSPSS 281

```

Search completed: March 23, 2006, 06:29:40
Job time : 10.911 secs

This Page Blank (usp)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: March 23, 2006, 04:56:29 ; Search time 31.904 Seconds
(without alignments)
1652.628 Million cell updates/sec

Title: US-10-617-038-30

Perfect score: 637
Sequence: 1 MSTRPRRHSGRVGPYAWA.....DMPAYAGHLSVETAVAV 120

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq_21.*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	637	100.0	120	8 ADI37309	Adi37309 M. tuberc
2	80	12.6	551	8 ADX68544	Adx68544 Plant ful
3	79.5	12.5	468	2 AAW71633	Aaw71633 Omega-CYC
4	78	12.2	310	7 ABO79202	AbO79202 Pseudomon
5	77.5	12.2	456	6 ABU40175	Abu40175 Protein e
6	75	11.8	510	7 ABO76512	AbO76512 Pseudomon
7	75	11.8	594	8 ADG22732	Adg22732 Cyanophag
8	75	11.8	8805	4 ABM67112	Abm67112 Drosophil
9	74.5	11.7	223	8 ADG65060	Adg65060 Novel hum
10	74	11.6	330	7 ABO77614	AbO77614 Pseudomon
11	74	11.6	535	6 ABU34179	Abu34179 Protein e
12	73	11.5	69	4 ABB39158	Abb39158 Peptide #
13	73	11.5	69	4 AAM32650	Aam32650 Peptide #
14	73	11.5	69	4 AAM72398	Aam72398 Human bon
15	73	11.5	69	4 AAM59811	Aam59811 Human bra
16	73	11.5	69	4 ABG54088	Abg54088 Human liv
17	73	11.5	69	5 ABG42216	Abg42216 Human pep
18	73	11.5	258	4 AAG75597	Aag75597 Human col
19	73	11.5	361	8 ADJ67604	Adj67604 Human ova
20	73	11.5	449	7 ADB64595	Adb64595 Human pro
21	73	11.5	715	4 AAU34857	Aau34857 E. coli c
22	73	11.5	715	6 ABU47919	Abu47919 Protein e
23	73	11.5	715	6 ABU15299	Abu15299 Protein e
24	73	11.5	715	8 ADS45230	Ads45230 Bacterial

25	73	11.5	734	4 AAB95634	Aab95634 Human pro
26	73	11.5	766	7 ADB79917	Adb79917 Human put
27	73	11.5	766	8 ABM81904	Abm81904 Tumour-as
28	73	11.5	766	8 ADS88275	Ads88275 Human pro
29	73	11.5	1277	4 ABB11784	Abb11784 Human sec
30	73	11.5	1277	8 ADS12057	Ads12057 Human the
31	73	11.5	1277	9 AEA20840	Aea20840 Novel hum
32	73	11.5	1277	9 AEA20840	Aea20840 Novel hum
33	72.5	11.4	221	6 AAU51687	Aau51687 Human ULP
34	72.5	11.4	221	6 ABM48206	Abm48206 Propionib
35	72.5	11.4	492	4 ABG14735	Abg14735 Novel hum
36	72	11.3	195	9 AEA79725	Aea79725 S. griseu
37	72	11.3	195	9 AEA48218	Aea48218 Streptogr
38	72	11.3	255	9 AEA80310	Aea80310 S. griseu
39	72	11.3	255	9 AEA48814	Aea48814 Mature st
40	72	11.3	457	9 AEA80324	Aea80324 Streptomy
41	72	11.3	457	9 AEA48827	Aea48827 Streptogr
42	72	11.3	510	2 AAR49835	Aar49835 Thermus a
43	71.5	11.2	759	6 ABE78096	Abp78096 N. gonorr
44	71	11.1	292	2 AAR88406	Aar88406 Trichoder
45	70.5	11.1	268	8 ADS73564	Ads73564 Glycopept

ALIGNMENTS

RESULT 1
ID ADI37309 standard; protein; 120 AA.
XX ADI37309;
AC
XX
DT 22-APR-2004 (first entry)
XX
DE M. tuberculosis low oxygen induced antigen Rv2628 SEQ ID NO:30.
XX
KW mycobacterial infection; vaccine; tuberculosis;
KW Mycobacterium tuberculosis; immunisation; antibacterial; gene therapy;
KW low oxygen induced antigen.
XX
OS Mycobacterium tuberculosis.
XX
PN WO2004006952-A2.
XX
PD 22-JAN-2004.
XX
PF 08-UTL-2003; 2003WO-DK000477.
XX
PR 13-UTL-2002; 2002DK-00001098.
XX
PA (STAT-) STATENS SERUM INST.
XX
PI Andersen P, Rosenkrands I, Stryhn A;
DR WPI; 2004-122778/12.
XX N-PSDB; ADI37354.
XX
PT Use of one or more polypeptides or their fragments, which are expressed
PT during the latent stage of the mycobacterial infection, and/or nucleic
PT acids encoding the polypeptides, for a therapeutic vaccine against
XX tuberculosis.
XX
PS Claim 3; SEQ ID NO 30; 76pp; English.
XX
CC The present invention describes polypeptides or their fragments, which
CC are expressed during the latent stage of a mycobacterial infection,
CC and/or nucleic acids encoding the polypeptides, which are useful for
CC creating a therapeutic vaccine against tuberculosis. Also described: (1)
CC a therapeutic vaccine against tuberculosis comprising one or more
CC polypeptides; (2) a method for treating an animal, including a human
CC being, with tuberculosis caused by virulent mycobacteria, e.g. by
CC Mycobacterium tuberculosis, M. africanum or M. bovis; (3) a method for
CC immunising an animal, including a human being, against tuberculosis

CC caused by virulent mycobacteria; (4) a method of diagnosing tuberculosis
CC caused by virulent mycobacteria in an animal, including a human being;
CC (5) a method for diagnosing previous or ongoing infection with a virulent
CC mycobacterium; and (6) a method of diagnosing mycobacterium tuberculosis
CC infection in a subject. The polypeptides have antibacterial activities,
CC and can be used in vaccines and in gene therapy. The polypeptides are
CC useful for the manufacture of a therapeutic vaccine for treating an
CC individual who is infected by a virulent mycobacterium, e.g. M.
CC tuberculosis, and who is not vaccinated with BCG against tuberculosis.
CC The present sequence represents a low oxygen induced antigen, which is
CC used in the exemplification of the present invention.

XX SQ Sequence 120 AA;

Query Match 100.0%; Score 637; DB 8; Length 120;
Best Local Similarity 100.0%; Pred. No. 5,1e-69;
Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGTQRRHSGIRAVGYAMAGRCGRIGRWGVHQAAMNNLAIMPRVQSATTYYQVDRSH 60
DB 1 MGTQRRHSGIRAVGYAMAGRCGRIGRWGVHQAAMNNLAIMPRVQSATTYYQVDRSH 60

QY 61 DGRITAVPQDEITSTVSGMLSELGTQSPLADELARAVRIGDMPAAYAGHLSVEIAVAV 120
DB 61 DGRITAVPQDEITSTVSGMLSELGTQSPLADELARAVRIGDMPAAYAGHLSVEIAVAV 120

RESULT 2
ADK68544
ID ADK68544 standard; protein; 551 AA.

XX AC ADK68544;

XX DT 21-APR-2005 (first entry)

XX DE Plant full length insert polypeptide seqid 39387.

KW plant protectant; plant growth regulant; gene therapy; plant;
KW recombinant DNA construct; physical array; plant breeding marker;
KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
KW extreme osmotic condition; pathogen tolerance; pest tolerance;
KW growth rate; cell cycle pathway; disease resistance;
KW galactomannan production; lignin production; plant growth regulator;
KW yield; plant growth; plant development; seed oil; protein yield;
KW protein content.

XX OS Undefined.

XX PN US2004034888-A1.

XX PD 19-FEB-2004.

XX PF 28-APR-2003; 2003US-0042514.

XX PR 06-MAY-1999; 99US-00304517.

XX PR 05-NOV-2001; 2001US-00985678.

XX PA (LIUJ/) LIU J.

XX PA (ZHOU/) ZHOU Y.

XX PA (KOVA/) KOVALIC D K.

XX PA (SCRE/) SCREEN S E.

XX PA (TABAS/) TABASKA J E.

XX PA (CAOY/) CAO Y.

XX PI Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;

XX DR WPI; 2004-180133/17.

XX PT New recombinant DNA construct, useful for improving plant tolerance to
XX cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
XX pests, for conferring increased resistance to plant disease, or for
XX improving yield.

PS Claim 1; SEQ ID NO 39387; 15pp; English.

XX The invention describes a recombinant DNA construct comprising a
XX polynucleotide consisting of a sequence encoding an amino acid sequence
XX available in electronic form from the US patent office at
XX ftp://seqdata.uspto.gov/sequence.html?docID:2004034888. The polynucleotide
XX of the invention are also useful in physical arrays of molecules and as
XX plant breeding markers. The recombinant DNA construct is useful for
XX improving plant tolerance to cold, heat, drought, herbicides, extreme
XX osmotic conditions, pathogens or pests, for manipulating growth rate in
XX plant cells by modification of the cell cycle pathway, for conferring
XX increased resistance to plant disease, for producing galactomannan,
XX lignin or plant growth regulators, for increasing the rate of homologous
XX recombination in plants, for improving yield by modification of
XX photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
XX or by providing improved plant growth and development under at least one
XX stress condition or for modifying seed oil or protein yield and/or
XX content. This is the amino acid sequence of a plant full length insert
XX polypeptide that can be used in the recombinant DNA construct of the
XX invention.

XX SQ Sequence 551 AA;

Query Match 12.6%; Score 80; DB 8; Length 551;
Best Local Similarity 27.0%; Pred. No. 1.9; Mismatches 43; Conservative 15; Indels 37; Gaps 10;

QY 6 PRH-----SGIRAVGYAMAGRCG-----RIGRWGVHQAAMNNLAIMPRKY----- 47
DB 71 PRHLPHGPRRRRRVRAPOPHL-HGRGGLREVVRAGEBQVNH--LHL-----PRVPLPR 123

QY 48 -----QSATTY-----QVDRSHDGTAVPQDEITSTVSGMLSELGTQSP 88
DB 124 PLGQHQLGRRRAGLYGMPRGHNLHQGTREHVSRRGLPRPR-----QLGSGVP 173

QY 89 LADELARAVRIGDMPAAYAGE-----HLVSVEIAVA 119
DB 174 QGANHARRAVG--PALVELGETEVAERPRHVGEBEYVA 210

RESULT 3
AAW71633
ID AAW71633 standard; protein; 468 AA.

XX AC AAW71633;

XX DT 26-NOV-1998 (first entry)

XX DE Omega-cyclohexane fatty acid biosynthesis enzyme #1 ORF1.

XX KW Omega-cyclohexane fatty acid; biosynthesis; enzyme; detection;

XX KW Alicyclobacillus acidocaldarius ATCC 27009; identification; microbe.

XX OS Alicyclobacillus acidocaldarius.

XX PN JP10234376-A.

XX PD 08-SEP-1998.

XX PF 28-FEB-1997; 97JP-00046570.

XX PR 28-FEB-1997; 97JP-00046570.

XX PA (KIRI) KIRIN BEVERAGE KK.

XX DR WPI; 1998-535030/46.

XX DR N-PSDB; AAV58229.

XX PT New nucleic acid - useful for detection and identification of genus
XX Alicyclobacillus microorganism(s).

XX PS Claim 5; Page 23-25; 37pp; Japanese.

CC The present sequence represents a protein from ORF1 of a new nucleic acid
CC which encodes enzymes which participate in the biosynthesis of Omega-
CC cyclohexane fatty acid. The nucleic acid is isolated from
CC *Allycyclobacillus acidocaldarius*. The present invention also describes
CC primers and probes containing all or part of the nucleic acid from
CC *Allycyclobacillus acidocaldarius*. The primers and probes may be used for
CC detection and/or identification of a microorganism of genus
CC *Allycyclobacillus*. The method can detect and identify *Allycyclobacillus*
CC genus rapidly and easily
XX
SQ Sequence 468 AA:

Query Match 12.5%; Score 79.5; DB 2; Length 468;
Best Local Similarity 28.3%; Pred. No. 1.8;
Matches 30; Conservative 11; Mismatches 40; Indels 25; Gaps 6;

QY 24 GRIGRGVGHQEAAMNLAIMHPRKVSATIVGVNDR-----SHDRTARVPEDETSTV--- 76
DB 368 GRSTRVAVRQSE-WORLGV--NVPVPMALDTQDGVRIERDRBELVFPADQVVLCTGQL 423
77 ---SGWLSEI-----GTQSPLADELARAVRIGDWPAAVAIG 109
DB 424 POSTATETKLPPIHVRVHIVGAGDSRDINARAIREA-WMAAYDIG 468

RESULT 4
ABO79202
ID ABO79202 standard; protein; 310 AA.
XX
AC ABO79202;
XX
DT 29-JUL-2004 (first entry)
XX
DE *Pseudomonas aeruginosa* polypeptide #11377.
XX
KM Bacterial infection; *Pseudomonas aeruginosa* infection; antibacterial.
XX
OS *Pseudomonas aeruginosa*.
XX
PN US651795-B1.
XX
PD 22-APR-2003.
XX
PF 18-FEB-1999; 99US-00252991.
XX
PR 18-FEB-1998; 98US-0074788P.
XX
PR 27-JUL-1998; 98US-0094190P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX
PI WPI; 2003-615309/58.
XX
DR N-PSDB; ABD12773.
XX
PT Novel isolated nucleic acid encoding *Pseudomonas aeruginosa* polypeptide,
XX useful as molecular targets for diagnostics, prophylaxis and treatment of
XX pathological conditions resulting from bacterial infection.
XX
PS Disclosure; SEQ ID NO 27948; 455bp; English.
XX
XX The invention relates to *Pseudomonas aeruginosa* polypeptides and the
XX polynucleotides encoding them. The sequences are useful in diagnosis and
XX therapy of pathological conditions, as molecular targets for diagnostics,
XX prophylaxis and treatment of pathological conditions resulting from a
XX bacterial infection, for evaluating a compound, such as a polypeptide,
XX for the ability to bind a *P. aeruginosa* nucleic acid, as components of
XX effective antibacterial targets, as targets for antibacterial drugs,
XX including anti-*P. aeruginosa* drugs, as templates for recombinant
XX production of *P. aeruginosa*-derived peptides or polypeptides, as target
XX components for diagnosis and/or treatment of *P. aeruginosa*-caused
XX infection, and in detection of *P. aeruginosa* sequences or other sequences
XX of *Pseudomonas* species using biochip technology. Sequences ABO67826-

CC ABO84396 represent *P. aeruginosa* polypeptides of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html
XX
SQ Sequence 310 AA:

Query Match 12.2%; Score 78; DB 7; Length 310;
Best Local Similarity 26.0%; Pred. No. 1.6;
Matches 27; Conservative 4; Mismatches 33; Indels 40; Gaps 3;

QY 9 SGIRAVGPYAMAGRCRIGRMKGVHQAAMNLAIMHPRKVSATIVGVNDRSHDGRPARVP 68
DB 237 AGRRCAGAGACGAGARRHG-----DGGTAKTA 263
69 GDEITSTVSGWLSELTQSPLADELARAVRIGDWPAAVAIGHL 112
DB 264 GERQAPPACTWVRQRLG---QCRQETARRRLRRQWPA-----GGHL 300

RESULT 5
ABU40175
ID ABU40175 standard; protein; 456 AA.
XX
AC ABU40175;
XX
DT 19-JUN-2003 (first entry)
XX
DE Protein encoded by prokaryotic essential gene #25702.
XX
KM Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
OS *Pseudomonas putida*.
XX
PN MO20027183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002WO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
XX
PR 06-SEP-2001; 2001US-00948993.
XX
PR 25-OCT-2001; 2001US-0342823P.
XX
PR 08-FEB-2002; 2002US-00072851.
XX
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zykkind JW;
XX
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
DR WPI; 2003-029926/02.
XX
DR N-PSDB; ACA44045.
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
XX for homologous nucleic acids required for cellular proliferation to
XX isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 25; SEQ ID NO 68099; 1766bp; English.
XX
XX The invention relates to an isolated nucleic acid comprising any one of
XX the 6213 antisense sequences given in the specification where expression
XX of the nucleic acid inhibits proliferation of a cell. Also included are:
XX (1) a vector comprising a promoter operably linked to the nucleic acid
XX encoding a polypeptide whose expression is inhibited by the antisense
XX nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX polypeptide or its fragment whose expression is inhibited by the
XX antisense nucleic acid; (4) an antibody capable of specifically binding
XX the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
XX proliferation or the activity of a gene in an operon required for
XX proliferation; (7) identifying a compound that influences the activity of
XX the gene product or that has an activity against a biological pathway
XX required for proliferation, or that inhibits cellular proliferation; (8)

CC for preparing enzymes for synthesis of D-bases (D = 2,6-diaminopurine),
 CC particularly D, dUMP and dTMP, or polynucleotides containing these bases,
 CC polymerases involved in metabolism of D-bases and deoxynucleotide
 CC analogs, for chemotherapy of AIDS. The genes, and encoded polypeptides,
 CC can be used for detection and/or identification of S-2L, and for
 CC identifying agents that modulate synthesis of D-bases or polynucleotides
 CC containing them, and fusions of S-2L polypeptides with an antigen can be
 CC used to raise specific antibodies, useful for detecting S-2L. This
 CC sequence corresponds to one of the proteins encoded by the cyanophage S-
 CC 2L genome.

XX Sequence 594 AA;

Query Match 11.8%; Score 75; DB 8; Length 594;

Best Local Similarity 31.2%; Pred. No. 8.7;
 Matches 25; Conservative 9; Mismatches 28; Indels 18; Gaps 4;

QY 58 RSHDGRTR-ARVPGDEITSTVSGWLSL-----GTGSPPLADELARAV-----RIGDM 102

DB 19 RHHGGRHQCQDGGEGERSIHFWAGSIALGHGTQAPIDSELPRIICPPQPADERLRIGQ 78

QY 103 PAAVAI---GEHLSVEIAVA 119

DB 79 PEAADAGRGPEHVLTVQVS 98

RESULT 8

ABB67112

ID ABB67112 standard; protein; 8805 AA.

AC ABB67112;

DT 26-MAR-2002 (first entry)

XX Drosophila melanogaster polypeptide SEQ ID NO 28128.

XX Drosophila; developmental biology; cell signalling; insecticide;

XX pharmaceutical.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US009231.

PR 23-MAR-2000; 2000US-0191637P.

PR 11-JUL-2000; 2000US-00614150.

XX (PEKE) PE CORP NY.

PI Venter JC, Adams M, Li PWD, Myers EW;

DR WPI; 2001-656860/75.

DR N-PSDB; ABL11215.

PT New isolated nucleic acid detection reagent for detecting 1000 or more

PT genes from Drosophila and for elucidating cell signalling and cell-cell

PT interactions.

PS Disclosure; SEQ ID NO 28128; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent

XX capable of detecting 1000 or more genes from Drosophila. The invention is

XX useful in developmental biology and in elucidating cell signalling and

XX cell-cell interactions in higher eukaryotes for the development of

XX insecticides, therapeutics and pharmaceutical drugs. The invention

XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA

XX sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-

XX ABB72072). The sequence data for this patent did not form part of the

XX printed specification, but was obtained in electronic format directly

XX from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 8805 AA;

Query Match 11.8%; Score 75; DB 4; Length 8805;

Best Local Similarity 29.9%; Pred. No. 2.9e+02;

Matches 29; Conservative 15; Mismatches 41; Indels 12; Gaps 4;

QY 28 RWGVHQAEMMNLAIWHPKRVQSAITVQYTDNR-HDGRFARV-----PDDEITSTVSGWL 80

DB 2326 RGLVDEQLIKIV--DPTTKQPISTVEAIARSISYDPKTIIDSEGPVDLITATKGLL 2383

QY 81 SELGTOSPLADELARAVRIGDMPAVAIGEHLVSVEIA 117

DB 2384 SVVGAFLVAABGALRTVRFVTDPRF---GEQIPVEVA 2417

AD065060 standard; protein; 223 AA.

AD065060;

DT 07-OCT-2004 (first entry)

DE Novel human protein sequence #33.

XX osteopathic; neuroprotective; nootropic; antiparkinsonian; cytostatic;

XX gene therapy; diagnostic marker; morbid state; osteoporosis;

XX neurological disease; Alzheimer's disease; Parkinson's disease; dementia;

XX cancer.

XX Homo sapiens.

XX EP1440981-A2.

XX 28-JUL-2004.

PF 21-JAN-2004; 2004EP-00001196.

PR 21-JAN-2003; 2003JP-00102206.

PR 09-MAY-2003; 2003JP-00131392.

XX (REAS-) RES ASSOC BIOTECHNOLOGY.

XX Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;

PI Yamamoto J, Isono Y, Nagai K, Irie R;

DR WPI; 2004-535376/52.

DR N-PSDB; ADQ62872.

PT Novel 2495 cDNA, useful for treating osteoporosis, neurological diseases,

PT Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.

XX Claim 1; SEQ ID NO 2221; 2449pp; English.

PS The invention relates to 2495 novel polynucleotides (I) and their encoded

XX polypeptides, sequences hybridizing to these nucleotides, sequences

XX encoding partial polypeptides and sequences having 70% or 90% identity to

XX the nucleotide and protein sequences. The nucleotides and polypeptides

XX are useful as diagnostic markers or therapeutic target for the diseases

XX or morbid states. They are also useful for treating osteoporosis,

XX neurological diseases, Alzheimer's diseases, Parkinson's diseases,

XX dementia and various cancers. This sequence corresponds to a protein

XX sequence of the invention.

XX Sequence 223 AA;

Query Match 11.7%; Score 74.5; DB 8; Length 223;

Best Local Similarity 27.3%; Pred. No. 2.8;

Matches 35; Conservative 13; Mismatches 29; Indels 51; Gaps 9;

QY 15 GPYVAGRGRIKRWGVHQAEMMNLAIW-----HPRKRVQSA-----TTIVQVT 56

DB 15 GPYVAGRGRIKRWGVHQAEMMNLAIW-----HPRKRVQSA-----TTIVQVT 56

Db 60 GWMALGR-RRIGR-----LALMAAPRRSGPRTTSEVGSRRPHRGFWRSREOS 108
QY 57 DRSHDR-TARVYGDDEITSTVSG--N--LSELTQSGPLA-----DELAR----- 95
Db 109 PPARGRCRGTVQVPGAGVSGVPGPTRWSAVPGCGERPFLARGRTAEGSGEPGRGVVPG 168
QY 96 -AVRIGDW 102
Db 169 AALRVGTW 176

RESULT 10
AB077614
ID AB077614 standard; protein; 330 AA.
XX
AC AB077614;
XX
DT 29-JUL-2004 (first entry)
XX
DE Pseudomonas aeruginosa polypeptide #9789.
XX
KM Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
XX
OS Pseudomonas aeruginosa.
XX
PN US651795-B1.
XX
PD 22-APR-2003.
XX
PF 18-FEB-1999; 99US-00252991.
XX
PR 18-FEB-1998; 98US-0074788P.
XX
PR 27-JUL-1998; 98US-0094190P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX
DR WPI; 2003-615309/58.
XX
DR N-PSDB; ABD11185.
XX
PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
XX
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
XX
PT pathological conditions resulting from bacterial infection.
XX
PS Disclosure; SEQ ID NO 26360; 455pp; English.
XX
XX The invention relates to Pseudomonas aeruginosa polypeptides and the
XX polynucleotides encoding them. The sequences are useful in diagnosis and
XX therapy of pathological conditions, as molecular targets for diagnostics,
XX prophylaxis and treatment of pathological conditions resulting from a
XX bacterial infection, for evaluating a compound, such as a polypeptide,
XX for the ability to bind a P. aeruginosa nucleic acid, as components of
XX effective antibacterial targets, as targets for antibacterial drugs,
XX including anti-P. aeruginosa drugs, as templates for recombinant
XX production of P. aeruginosa-derived peptides or polypeptides, as target
XX components for diagnosis and/or treatment of P. aeruginosa-caused
XX infection, and in detection of P. aeruginosa sequences or other sequences
XX of Pseudomonas species using biochip technology. Sequences AB067826-
XX AB084396 represent P. aeruginosa polypeptides of the invention. Note: The
XX sequence data for this patent did not form part of the printed
XX specification but was obtained in electronic format from USPTO at
XX seqdata.uspto.gov/sequence.html
SQ Sequence 330 AA;

Query Match 11.6%; Score 74; DB 7; Length 330;
Best Local Similarity 25.2%; Pred. No. 5.3;
Matches 31; Conservative 13; Mismatches 59; Indels 20; Gaps 4;
QY 2 SGRGRPHSGIRAVGPAAW-----GRCGRIGW-----GHOEAMMNLAIWH 43
Db 26 TSRRRPGPLPMAAPVMAARRRRLGSGRSBSRSGRWPSRVCTTSMPTRAOSSSAMIATGA 85

QY 44 PRKVSATITYQVTDGS-HDGR TARVPGDEITSTVSGWLTSELTQSGPLADELARAVRIGDW 102
Db 86 RRLRTISIPAMSAVYPPSAQSRCSITSAVCSRSSGCVRLTAQPTCPR-RRCLRRCRG 144
QY 103 PAA 105
Db 145 PAA 147

RESULT 11
ABU34179
ID ABU34179 standard; protein; 535 AA.
XX
AC ABU34179;
XX
DT 19-JUN-2003 (first entry)
XX
DE Protein encoded by prokaryotic essential gene #19706.
XX
KM Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
OS Mycobacterium avium.
XX
PN WO20027183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002MO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
XX
PR 06-SEP-2001; 2001US-00948993.
XX
PR 25-OCT-2001; 2001US-0342923P.
XX
PR 08-FEB-2002; 2002US-00072851.
XX
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.
PI Wang L, Zamudio C, Malone C, Haselebeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Foreys RA, Xu HH;
XX
XX WPI; 2003-029926/02.
XX
DR N-PSDB; ACA38049.
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
XX
PT for homologous nucleic acids required for cellular proliferation to
XX
PT isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 25; SEQ ID NO 62103; 1766pp; English.
XX
XX The invention relates to an isolated nucleic acid comprising any one of
XX the 6213 antisense sequences given in the specification where expression
XX of the nucleic acid inhibits proliferation of a cell. Also included are:
XX (1) a vector comprising a promoter operably linked to the nucleic acid
XX encoding a polypeptide whose expression is inhibited by the antisense
XX nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX polypeptide or its fragment whose expression is inhibited by the
XX antisense nucleic acid; (4) an antibody capable of specifically binding
XX the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
XX proliferation or the activity of a gene in an operon required for
XX proliferation; (7) identifying a compound that influences the activity of
XX the gene product or that has an activity against a biological pathway
XX required for proliferation, or that inhibits cellular proliferation; (8)
XX identifying a gene required for cellular proliferation or the biological
XX pathway in which a proliferation-required gene or its gene product lies
XX or a gene on which the test compound that inhibits proliferation of an
XX organism acts; (9) manufacturing an antibiotic; (10) profiling a
XX compound's activity; (11) a culture comprising strains in which the gene
XX product is overexpressed or underexpressed; (12) determining the extent
XX to which each of the strains is present in a culture or collection of
XX strains; or (13) identifying the target of a compound that inhibits the
XX proliferation of an organism. The antisense nucleic acids are useful for
XX identifying proteins or screening for homologous nucleic acids required


```
Db          2 ELSDVSGGSKATP-----ASTANSDVATIPDTPLKENEGRVKTADAPNKSEISKHIE 57
Qy          114 VEIA 117
           |::|
Db          58 VQVA 61

RESULT 14
AAM72398
ID AAM72398 standard; protein; 69 AA.
AC AAM72398;
DT 06-NOV-2001 (first entry)
XX
DE Human bone marrow expressed probe encoded protein SEQ ID NO: 32704.
XX
KW Human; bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; Leukemia; Lymphoma; myeloma.
OS Homo sapiens.
PN MO200157276-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000668.
XX
PR 04-FEB-2000; 2000US-0180312P.
XX
PR 26-MAY-2000; 2000US-0207456P.
XX
PR 30-JUN-2000; 2000US-00608408.
XX
PR 03-AUG-2000; 2000US-00632366.
XX
PR 21-SEP-2000; 2000US-0234687P.
XX
PR 27-SEP-2000; 2000US-0236359P.
XX
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488900/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human bone marrow.
XX
PS Example 4; SEQ ID NO 32704; 658bp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukemia and myeloma. The present sequence is a
CC protein encoded by one of the probes of the invention
XX
SQ Sequence 69 AA;

Query Match          11.5%; Score 73; DB 4; Length 69;
Best Local Similarity 28.1%; Pred. No. 0.92;
Matches 18; Conservative 15; Mismatches 27; Indels 4; Gaps 1;

Qy          54 QVTRSHDGRTRAVPDEITSTVSGWLSGLTQSPDLADLARAVRIGDMPAAYAIGEHLIS 113
           |::|
Db          2 ELSDVSGGSKATP-----ASTANSDVATIPDTPLKENEGRVKTADAPNKSEISKHIE 57

Qy          114 VEIA 117
           |::|
Db          58 VQVA 61

RESULT 15
AAM59811
ID AAM59811 standard; protein; 69 AA.
```

```
AC AAM59811;
DT 05-NOV-2001 (first entry)
XX
DE Human brain expressed single exon probe encoded protein SEQ ID NO: 31916.
XX
KW Human; brain expressed exon; gene expression analysis; probe; microarray;
KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer.
XX
OS Homo sapiens.
PN MO200157275-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000667.
XX
PR 04-FEB-2000; 2000US-0180312P.
XX
PR 26-MAY-2000; 2000US-0207456P.
XX
PR 30-JUN-2000; 2000US-00608408.
XX
PR 03-AUG-2000; 2000US-00632366.
XX
PR 21-SEP-2000; 2000US-0234687P.
XX
PR 27-SEP-2000; 2000US-0236359P.
XX
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483446/52.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT brains.
XX
PS Example 4; SEQ ID NO 31916; 650bp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancer. The present sequence is a protein encoded by one of
CC the probes of the invention
XX
SQ Sequence 69 AA;

Query Match          11.5%; Score 73; DB 4; Length 69;
Best Local Similarity 28.1%; Pred. No. 0.92;
Matches 18; Conservative 15; Mismatches 27; Indels 4; Gaps 1;

Qy          54 QVTRSHDGRTRAVPDEITSTVSGWLSGLTQSPDLADLARAVRIGDMPAAYAIGEHLIS 113
           |::|
Db          2 ELSDVSGGSKATP-----ASTANSDVATIPDTPLKENEGRVKTADAPNKSEISKHIE 57

Qy          114 VEIA 117
           |::|
Db          58 VQVA 61

Search completed: March 23, 2006, 05:10:58
Job time : 33.904 secs
```


A:Residues: 1-691 <TEXT>
A:Cross-references: UNIPROT:Q9K0B2; UNIPARC:UPI00000C4543; GB:AE002425; GB:AE002098; NITD
A:Experimental source: serogroup B, strain KC58
C:Genetics:
A:Gene: NMB0702
C:Superfamily: competence protein ComEC

Query Match 13.14%; Score 85.5; DB 2; Length 691;
Best Local Similarity 27.1%; Pred. No. 1.1;
Matches 29; Conservative 14; Mismatches 51; Indels 13; Gaps 3;

OY 6 PRHSGIRAVGPYAWAGRCGRIGRWGVHOEAMNNTAIWPRKY-----OSATIQ 54
Db PTGSRVTLAAAFAMWMRRGRLSAMATWWQALAAVLFPPLAVLGVTGLSRGLVAALIWA 300
: : | | | | : : : : : : : : : :
55 VTRSDHG-RTARVPGEDEITSTVS-GWLSELTGSPLADELARAVRI 99
: : : : : : : : : : : : : : : :
301 CSGRLEHGKRQTALRGQMAASVLSVLILGYFASFLPSLIVNAVAI 347

RESULT 3
BB1937
competence protein MMA0906 [imported] - Neisseria meningitidis (strain Z2491 serogroup A)
C:Species: Neisseria meningitidis
C:date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 09-Jul-2004
C:Accession: BB1937
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel,
J.; Holtz, S.; Jogle, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,
A.Tille: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.
A:Reference number: AB1775; MUID:20222556; PMID:10761919
A:Accession: BB1937
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-691 <PAR>
A:Cross-references: UNIPROT:O9JVB8; UNIPARC:UPI00000C4A92; GB:AL162754; GB:AL157959; NITD
A:Experimental source: serogroup A, strain Z2491
C:Genetics:
A:Gene: comA; MMA0906
C:Superfamily: competence protein ComEC

Query Match 13.1%; Score 83.5; DB 2; Length 691;
Best Local Similarity 26.2%; Pred. No. 1.8;
Matches 28; Conservative 15; Mismatches 51; Indels 13; Gaps 3;

OY 6 PRHSGIRAVGPYAWAGRCGRIGRWGVHOEAMNNTAIWPRKY-----OSATIQ 54
Db PTGSRVTLAAAFAMWMRRGRLSAMATWWQALAAVLFPPLAVLGVTGLSRGLVAALIWA 300
: : | | | | : : : : : : : : : :
55 VTRSDHG-RTARVPGEDEITSTVS-GWLSELTGSPLADELARAVRI 99
: : : : : : : : : : : : : : : :
301 CSGRLEHGKRQTALRGQMAASVLSVLILGYFASFLPSLIVNAVAI 347

RESULT 4
GG8145
probable aminotransferase y4ub [imported] - Agrobacterium tumefaciens (strain C58, Cerec
C:Species: Agrobacterium tumefaciens
C:date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 31-Dec-2004
C:Accession: GG8145
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quirollo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A>Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A:Reference number: AF7359; MUID:21608551; PMID:11743194
A:Accession: GG8145
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-520 <KUD>
A:Cross-references: UNIPROT:Q8U6P6; UNIPARC:UPI00000D26BF; GB:AE007670; PIDN:AAK68689.1;
C:Genetics:
A:Gene: AGR_L_242
A:Map position: linear chromosome

[illegible]

A:Accession: G87496
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-609 <STO>
A:Cross-references: UNIPROT:Q9NA6T7; UNIPARC:UPI00000C758F; GB:AEO05673; NID:G13423465; F
C:Genetic8:
A:Gene: CCl1996

Query Match 11.6%; Score 74; DB 2; Length 609;
Best Local Similarity 26.0%; Pred. No. 14;
Matches 26; Conservative 9; Mismatches 39; Indels 26; Gaps 3;

Oy 15 GPYAMAGRCGRIGRMVGHQAMNMLATMHPKVSATITQVTDNRSDRTARVP-GDETT 73
Db 480 GRPFWLRKDIIISFGV-----KGIGQRNDGLNTAPGGTFPLV 517

Oy 74 STVGWLSIEGTOSPLADELARAVRIGDMPAAVAIGEHLIS 113
Db 518 SSADGEIAYAGNQVPTFGNLVLVGHADGWTAIV---HLS 554

RESULT 7
DEFCS
formate dehydrogenase (EC 1.2.1.2) H (hydrogenase-linked) [validated] - Escherichia coli
N:Alternate names: formate dehydrogenase H (benzylviologen-linked) (PDH-H); formate hyd
C:Species: Escherichia coli
C:Date: 31-Mar-1988 #basecount, revision 31-Dec-1992 #text_change 05-Oct-2004
C:Accession: A24145; F65216; A36088
R:Zinnion, F.; Birkmann, A.; Stadman, T.C.; Boeck, A.
Proc. Natl. Acad. Sci. U.S.A. 83, 4650-4654, 1986
A>Title: Nucleotide sequence and expression of the selenocysteine-containing polypeptide
A:Reference number: A24145; MUID:86259655; PMID:2941757
A:Accession: A24145
A:Molecule type: DNA
A:Residues: 1-139,'X','141-715 <ZIN>
A:Cross-references: UNIPROT:P07658; UNIPARC:UPI000011043F; GB:M13563; GB:M18632; NID:G14
A:Experimental source: strain MC4100
A>Note: The authors translated the selenocysteine UGA codon for residue 140 as 'X'; we h
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
A.: Rose, D.J.; Mau, B.; Shaoh, Y.
Science 277, 1453-1462, 1997
A>Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: F65216
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-139,'X','141-715 <BLAT>
A:Cross-references: UNIPARC:UPI000011043F; GB:AEO00481; GB:U00096; NID:G2367346; PIDN:AA
A:Experimental source: strain K-12, Substrain MC1655
A>Note: the authors translated the selenocysteine UGA codon for residue 140 as 'X'; we h
R:Axley, M.D.; Grahame, D.A.; Stadman, T.C.
J. Biol. Chem. 265, 18213-18218, 1990
A>Title: Escherichia coli formate-hydrogen lyase. Purification and properties of the sel
A:Reference number: A36088; MUID:91009305; PMID:2211698
A:Accession: A36088
A:Molecule type: protein
A:Residues: 1-7,'X','9-10,'X','12 <AXL>
A:Cross-references: UNIPARC:UPI0000171FB3
R:Gladychev, V.N.; Khangulov, S.V.; Axley, M.J.; Stadman, T.C.
Proc. Natl. Acad. Sci. U.S.A. 91, 7708-7711, 1994
A>Title: Coordination of selenium to molybdenum in formate dehydrogenase H from Escheric
A:Reference number: A58332; MUID:94329585; PMID:8052647
A:Contents: annotation; selenium coordination of cofactor by EPR
R:Sun, P.D.; Boyington, J.C.
submitted to the Brookhaven Protein Data Bank, January 1997
A:Reference number: A67458; PDB:1AB6
A:Contents: annotation; X-ray crystallography, 2.3 angstroms, reduced form, residues 1-6
R:Sun, P.D.; Boyington, J.C.
submitted to the Brookhaven Protein Data Bank, January 1997
A:Reference number: A67589; PDB:1FDI
A:Contents: annotation; X-ray crystallography, 2.9 angstroms, oxidized form, residues 1-
R:Sun, P.D.; Boyington, J.C.
submitted to the Brookhaven Protein Data Bank, January 1997

[illegible]

Db 666 EPIADORAALQYVIDEY 702

RESULT 9

JC5958

sperm surface protein - human

C:Species: Homo sapiens (man)

C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004

C:Accession: JC5958

R:Shankar, S.; Mohapatra, B.; Suri, A.

Biochem. Biophys. Res. Commun. 243, 561-565, 1998

A:Title: Cloning of a novel human testis mRNA specifically expressed in testicular haplo

A:Reference number: JC5958; MUID:98153804; PMID:9480848

A:Accession: JC5958

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-766 <SHA>

A:Cross-references: UNIPROT:O60905; UNIPARC:UPI000017C354; EMBL:X91879

Query Match 11.5%; Score 73; DB 2; Length 766;
Best Local Similarity 28.1%; Pred. No. 23;

Matches 18; Conservative 15; Mismatches 27; Indels 4; Gaps 1;

QY 54 QVTDRSHDRTARVPDEITSTVSGWLSBLGQSPDLADLARAARIGDWPAAVAIGEHLS 113

Db 84 ELSDVSGGSKATTP---ASTANSVATIPDTPLKEENEGFVKVTDAPNKEISKHLE 139

QY 114 VEIA 117

Db 140 VOVA 143

RESULT 10

SS7953

C4BP protein alpha chain precursor - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 13-Jan-1996 #sequence_revision 01-Mar-1996 #text_change 09-Jul-2004

C:Accession: SS7953

R:Hillarp, A.; Thern, A.; Dahlback, B.

submitted to the EMBL Data Library, July 1995

A:Title: Molecular cloning of rat C4b-binding protein alpha- and beta-chains: stru

A:Reference number: SS7953

A:Accession: SS7953

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-558 <HL>

A:Cross-references: UNIPROT:O63514; UNIPARC:UPI0000126C29; EMBL:Z50051; NID:G899379; PID

C:Superfamily: C4b-binding protein alpha chain; complement factor H repeat homology

C:Comment: This enzyme hydrolyzes the peptide bond at the C-terminus of peptides and pro

C:Keywords: This enzyme is a metallo-carboxypeptidase with thermostability and a broad su

A:Gene: Tag

A:Superfamily: Zn-dependent carboxypeptidase, Tag type

C:Keywords: hydrolase

Query Match 11.4%; Score 72.5; DB 2; Length 558;
Best Local Similarity 29.9%; Pred. No. 18;

Matches 26; Conservative 11; Mismatches 37; Indels 13; Gaps 4;

QY 23 CGRIGRW-----GVHQAAMNLAIWPRKVSATITYT-----DRSHDGRFARVPGDEITS 74

Db 373 CKSDGTHWRTPSCHQSCDFPALAHGRYTKSSVYRVQVTECEBG--YLVGE---A 427

QY 75 TVSGWLSBLGQSPDLADLARAARIGD 101

Db 428 TISCWYSQWTPAAPQCKALCRKEIGN 454

RESULT 11

A53669

streptogrisin C (EC 3.4.21.-) precursor - Streptomyces griseus

C:Species: Streptomyces griseus

C:Date: 07-Oct-1994 #sequence_revision 07-Oct-1994 #text_change 05-Oct-2004

C:Accession: A53669

R:Sidhu, S.S.; Kaimar, G.B.; Willis, L.G.; Borgford, T.J.

J. Biol. Chem. 269, 20167-20171, 1994

A:Title: Streptomyces griseus protease C. A novel enzyme of the chymotrypsin superfamily

A:Reference number: A53669; MUID:94327573; PMID:8051104

A:Accession: A53669

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-457 <SID>

A:Cross-references: UNIPROT:P52320; UNIPARC:UPI00000139F; GB:L29018; NID:G532654; PIDN...

C:Species: Streptomyces griseus

C:Date: 07-Oct-1994 #sequence_revision 07-Oct-1994 #text_change 05-Oct-2004

C:Accession: A53669

R:Sidhu, S.S.; Kaimar, G.B.; Willis, L.G.; Borgford, T.J.

J. Biol. Chem. 269, 20167-20171, 1994

A:Title: Streptomyces griseus protease C. A novel enzyme of the chymotrypsin superfamily

A:Reference number: A53669; MUID:94327573; PMID:8051104

A:Accession: A53669

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-457 <SID>

A:Cross-references: UNIPROT:P52320; UNIPARC:UPI00000139F; GB:L29018; NID:G532654; PIDN...

C:Genetics:

A:Start codon: GTG

C:Superfamily: serine protease precursor, streptogrisin type

C:Keywords: hydrolase, serine proteinase

Query Match 11.3%; Score 72; DB 2; Length 457;
Best Local Similarity 28.0%; Pred. No. 16;

Matches 26; Conservative 14; Mismatches 27; Indels 26; Gaps 5;

QY 17 YAMAGRCRIGRW--GVHQA-----MMNLAIWPRKVSATITYQVDR 58

Db 233 FATAGHGRVGTGTTNGVQAQGTFGSTFPGRDIAMVATNMTPRLVNG--YGRBDV 290

QY 59 SHDGRFARVPGDEI--TSTVSGW---LSEIGT 85

Db 291 TVAGSTASVAGSVCRSGSTTGWHGRIQOLNT 323

RESULT 12

JC2294

carboxypeptidase (EC 3.4.1.-) Tag - Thermus aquaticus (strain YT-1)

C:Species: Thermus aquaticus

C:Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 05-Oct-2004

C:Accession: JC2294; PC2202

R:Lee, S.H.; Tsuchi, H.; Yoshimura, E.; Mingawa, E.; Kamimura, S.; Ohta, T.; Matsumura

Biosci. Biotechnol. Biochem. 58, 1490-1495, 1994

A:Title: Carboxypeptidase Tag, a thermostable zinc enzyme, from Thermus aquaticus YT-1:

A:Reference number: JC2294; MUID:94369103; PMID:7765282

A:Accession: JC2294

A:Molecule type: DNA

A:Residues: 1-511 <LEB1>

A:Cross-references: UNIPROT:P42663; UNIPARC:UPI00001285FD; DDBJ:D17669; NID:G497631; PID:

A:Accession: PC2202

A:Molecule type: protein

A:Residues: 1-511 <LEB2>

A:Cross-references: UNIPARC:UPI00001285FD

C:Comment: This enzyme hydrolyzes the peptide bond at the C-terminus of peptides and pro

C:Keywords: This enzyme is a metallo-carboxypeptidase with thermostability and a broad su

A:Gene: Tag

A:Superfamily: Zn-dependent carboxypeptidase, Tag type

C:Keywords: hydrolase

Query Match 11.3%; Score 72; DB 2; Length 511;
Best Local Similarity 27.2%; Pred. No. 18;

Matches 31; Conservative 14; Mismatches 37; Indels 32; Gaps 6;

QY 18 AMAGRCRIGRWGVHQAAMNLAIWPRKVSATITYQT-----DRSHDGRFARVPGD 70

Db 18 AYLSGLGLAW--DQRTM-----PRKGHGRARQVALARLLHERATPR----- 62

QY 71 EITSTVSGWLSBLGQSPDLADLARAARIGDWPAAV---AIGEHLSVEIYVA 119

Db 63 -----IGEWLEKVGSSLVEDPLDPAVNVAMRRARARARAVELAVELAA 111

RESULT 13

G70932

probable monooxygenase cytochrome P450 Rv0568 - Mycobacterium tuberculosis (strain H37Rv)

N:Contains: oxidoreductase (EC 1.-.-.-)

C:Species: Mycobacterium tuberculosis

C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
 C/Accession: G70932
 R/Collection: S.T.; Broesch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Cole, S.T.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holtroyd, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holtroyd, S.; Rajandram, M.A.; Rogers, J.; Ruter, S.; Seeger, K.; Skelton, S.; Squares, S.
 Nature 393, 537-544, 1998
 A/Authors: Bares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A/Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome
 A/Reference number: A70500; PMID:98295987; PMID:9634230
 A/Accession: G70932
 A/Status: preliminary; nucleic acid sequence not shown; translation not shown
 A/Molecule type: DNA
 A/Residues: 1-472 <COL>
 A/Cross-references: UNIPROT:O53765; UNIPARC:UP10000126B9; GB:AL021942; GB:AL123456; NID
 A/Experimental source: strain H37Rv
 C/Genetics:
 A/Genes: RV0568
 C/Superfamily: Bacillus cytochrome P450 CYP106; cytochrome P450 homology
 C/Keywords: chromoprotein; heme; iron; metalloprotein; oxidoreductase
 F/258-410/Domain: cytochrome P450 homology <P45>
 F/388/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 11.1%; Score 71; DB 2; Length 472;
 Best Local Similarity 21.0%; Pred. No. 21;
 Matches 37; Conservative 23; Mismatches 46; Indels 70; Gaps 9;

QY 7 RRSGRVAVPVYAWGRGCRIGRW---GVHQAAMNLAH--PRKQ--SATY--QVTR 58
 DB 165 RLAAALRKWP-----RLAVGPMATLALNPISLNRRLSRRLREEDALLVETADR 219
 QY 59 -----SHDGRAR-----VPGDEITSTVSGMLSELGTQ 86
 DB 220 RADPLAARTDTLAMLVRAADEGRMTRELRDQLITLLVAGHDTTATGLSWALERLTR 279

QY 87 SPLADELAAYVIGD-----WPAAYAGEHLSEIAYV 119
 DB 280 HPVT--LAKAVQAAADASAAGDPAGDEYLDVAKETLRIRPVYDVGRVLTAEVVA 333

RESULT 14
 AF0666
 Probable exported protein STY1442 [imported] - *Salmonella enterica* subsp. *enterica* serov. *C*
 C/Species: *Salmonella enterica* subsp. *enterica* serovar Typh
 A/Note: this species has also been called *Salmonella typhi*
 C/Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
 C/Accession: AF0666
 R/Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, T.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holtroyd, S.; Rajandram, M.A.; Rogers, J.; Ruter, S.; Seeger, K.; Skelton, S.; Squares, S.; S.; Moule, S.; O'Gaora, P.
 Nature 413, 848-852, 2001
 A/Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; A/Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* serov. *A*
 A/Reference number: AB0502; PMID:11677608
 A/Accession: AF0666
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-541 <PAR>
 A/Cross-references: UNIPARC:UP10000059EB8; GB:AL513382; PIDN:CAD01703.1; PID:G16502553; C/Genetics:
 A/Genes: STY1442
 C/Superfamily: periplasmic glucans biosynthesis protein mdog

Query Match 11.1%; Score 71; DB 2; Length 541;
 Best Local Similarity 22.9%; Pred. No. 24;
 Matches 22; Conservative 11; Mismatches 33; Indels 30; Gaps 3;

QY 28 RMGVQAEAMNML-----AIWPKVQSAITYQVDRSHDGRARVPGDEITSTV 76
 DB 359 KMGKAVSLMEIPTGETIDNVVCFWQPRKAIA-----GDTLAFNY 400

QY 77 S-GWLSLGTGSPPLADELARAVRIGDMPAAVAGEH 111
 DB 401 RLYWSAQPVPVQSPPLARVMAATRTMGGFPEGMAFGEH 436

RESULT 15
 S75260
 Hypothetical protein gl0985 - *Synechocystis* sp. (strain PCC 6803)
 C/Species: *Synechocystis* sp.
 A/Variety: PCC 6803
 C/Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
 C/Accession: S75260
 R/Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.; O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
 DNA Res. 3, 109-136, 1996
 A/Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp.
 A/Reference number: S74322; PMID:97061201; PMID:8905231
 A/Accession: S75260
 A/Status: nucleic acid sequence not shown; translation not shown
 A/Molecule type: DNA
 A/Residues: 1-704 <KAN>
 A/Cross-references: UNIPROT:P73148; UNIPARC:UP100000C10A9; EMBL:D90904; GB:AB001339; NID
 A/Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
 C/Genetics:
 A/Start codon: GTG

Query Match 11.1%; Score 71; DB 2; Length 704;
 Best Local Similarity 29.0%; Pred. No. 33;
 Matches 20; Conservative 10; Mismatches 27; Indels 12; Gaps 2;

QY 42 WPKVQSAITYQVDRSHDGRARVPGDEITSTVSGMLSEL-GTQSPPLADELARVRIG 100
 DB 617 WPKVQSAITYQVDRSHDGRARVPGDEITSTVSGMLSEL-GTQSPPLADELARVRIG 100

QY 101 DWPAAYAG 109
 DB 666 PESQETALG 674

Search completed: March 23, 2006, 05:26:37
 Job time : 5.77423 secs

THIS PAGE BLANK (00710)

This page Blank (uspric)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: March 23, 2006, 04:56:43 ; Search time 30.6915 Seconds

(without alignments)
2758.529 Million cell updates/sec

Title: US-10-617-038-30

Perfect score: 637
Sequence: 1 MSTOPRRHSRGIRAVGPYMA.....DWPAAYALGHLSEIVAVAV 120

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot_05.80:*

1: uniprot_sprot:*

2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	637	100.0	120	006184_MYCTU	006184 mycobacteri
2	622	97.6	120	077Y70_MYCBO	077Y70 mycobacteri
3	260	40.8	87	073MD2_MYCBO	073MD2 mycobacteri
4	87.5	13.7	447	051LJ1_SILPO	051LJ1 silicibacte
5	85.5	13.4	691	09KOB2_NEIMB	09KOB2 neisseria m
6	83.5	13.1	691	09JVB8_NEIMA	09JVB8 neisseria m
7	82.5	13.0	331	05YS68_NOCFA	05YS68 nocardia fa
8	82	12.9	281	07VVV9_BORBP	07VVV9 bordetella
9	81	12.7	277	07WKL0_BORBR	07WKL0 bordetella
10	80.5	12.6	263	08ZT55_PYPAB	08ZT55 pyrobaculum
11	80.5	12.6	520	07CCT0_AGR15	07CCT0 agrobacteri
12	79.5	12.5	691	1 COMA_NETGO	P51973 neisseria g
13	79	12.4	270	05KXV7_GROKA	05KXV7 geobacillus
14	78	12.2	163	093MF2_AMYMD	093MF2 amycolatops
15	77.5	12.2	456	088E24_PSRBP	088E24 pseudomonas
16	77	12.1	277	07N269_CHRVO	07N269 chromobace
17	77	12.1	316	07W2W1_BORPA	07W2W1 bordetella
18	76.5	12.0	283	04KE05_PSRFS	04KE05 pseudomonas
19	76	11.9	455	04K5Q9_PSRFS	04K5Q9 pseudomonas
20	76	11.9	757	05GZM7_XANOR	05GZM7 xanthomonas
21	75.5	11.9	316	07WDM1_BORBR	07WDM1 bordetella
22	75.5	11.9	510	05SLM3_THRT8	05SLM3 thermus the
23	75.5	11.9	510	07Z6Y3_THRT8	07Z6Y3 thermus the
24	75	11.8	337	04RMT2_THRT2	04RMT2 tetradion n
25	75	11.7	8805	07KR73_DROME	07KR73 drosophila
26	74.5	11.7	223	06ZMG9_HUMAN	06ZMG9 homo sapien
27	74.5	11.7	252	05Z1Y0_NOCFA	05Z1Y0 nocardia fa
28	74.5	11.7	527	0885N9_PSRSM	0885N9 pseudomonas
29	74.5	11.7	731	08FMS1_CORBF	08FMS1 corynebacte
30	74	11.6	284	0820D7_STROPM	0820D7 streptomyce
31	74	11.6	609	09A6T7_CAUCR	09A6T7 caulobacter

32	73.5	11.5	214	2	05U2U1_RAT	05U2U1 rattus norv
33	73.5	11.5	223	2	067U16_ORYSA	067U16 oryza sativ
34	73.5	11.5	216	2	054MF6_DICDI	054MF6 dictyosteli
35	73.5	11.5	407	2	09KYV1_STRCO	09KYV1 streptomyce
36	73	11.5	173	2	0889T1_PSRSM	0889T1 pseudomonas
37	73	11.5	204	2	066WC7_HUMAN	066WC7 homo sapien
38	73	11.5	365	2	08FH77_ECOL6	08FH77 escherichia
39	73	11.5	712	2	070J33_MACFA	070J33 macaca fasc
40	73	11.5	715	1	PDHF_ECOLI	P07658 escherichia
41	73	11.5	715	2	05PJ14_SALPA	05PJ14 salmonella
42	73	11.5	715	2	08PAX4_ECOL6	08PAX4 escherichia
43	73	11.5	715	2	08ZKE7_SALTY	08ZKE7 salmonella
44	73	11.5	715	2	08Z1U0_SALTY	08Z1U0 salmonella
45	73	11.5	734	2	09H811_HUMAN	09H811 homo sapien

ALIGNMENTS

```
RESULT 1
006184 MYCTU PRELIMINARY; PRT; 120 AA.
ID 006184 MYCTU PRELIMINARY; PRT; 120 AA.
AC 006184 07D6V2;
DT 01-JUL-1997 (TRMBLrel. 04, Created)
DT 01-JUL-1997 (TRMBLrel. 04, Last sequence update)
DT 13-SEP-2005 (TRMBLrel. 31, Last annotation update)
DE Hypochemical protein.
GN OrderedAccessionNames=MT2703, RV2628;
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinomycetales;
OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium tuberculosis complex.
OX NCBI_TaxID=1773;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230; DOI=10.1038/31159;
RA Cole S.T., Brosch R., Parkhill J., Garfield T., Churcher C.M.,
RA Harris D.E., Gordon S.V., Eigilmeier K., Gas S., Barry C.E. III,
RA Tekala F., Badcock K., Basham D., Brown D., Chillingworth T.,
RA Connor R., Davies R.M., Devlin K., Felwell T., Gentles S., Hamlin N.,
RA Holtroyd S., Hornbly T., Jagels K., Krogh A., McLean J., Moule S.,
RA Murphy L.D., Oliver S., Osborne J., Quail M.A., Rajandream M.A.,
RA Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulten J.E., Taylor K., Whitehead S., Barrett B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544 (1998).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CDC 1551 / Oshkosh;
RX MEDLINE=22206494; PubMed=12218035;
RA DOI=10.1128/JB.184.19.5479-5490.2002;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J.D., DeBoy R.T., Dodson R.J., Gwinn M.L., Haft D.H.,
RA Hickey E.K., Kolony J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D.,
RA Salberg S.L., Delcher A., Uitterback T.R., Weidman J.F., Khouri H.M.,
RA Gill J., Mikula A., Bishai W., Jacobs W.R. Jr., Venter J.C.,
RA Fraser C.M.;
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL J. Bacteriol. 184:5479-5490 (2002).
DR EMBL; BX842580; CAB08635.1; -; Genomic DNA.
DR EMBL; AE000515; AAK47019.1; -; Genomic DNA.
DR PIR; C70573; C70573.
DR TIGR; MT2703; -.
DR TubercleList; RV2628; -.
KW Complete proteome; Hypochemical protein.
SQ SEQUENCE 120 AA; 13129 MM; CAAAG642A61B42F0 CRC64;
Query Match 100.0%; Score 637; DB 2; Length 120;
Best Local Similarity 100.0%; Pred. No. 2; Le-55;
Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 MSTOPRRHSGIRAVGYAMAGRCGRIGRWGVHGOEAMNNLAIWHPKRVGSATTIYQVTDNRSH 60
Db 1 MSTOPRRHSGIRAVGYAMAGRCGRIGRWGVHGOEAMNNLAIWHPKRVGSATTIYQVTDNRSH 60
Qy 61 DQRTARVPDEITSTVSGWLSLGTQSPPLADELARAIVRIGDMPAAVAIGEHLVSEIYAVAV 120
Db 61 DQRTARVPDEITSTVSGWLSLGTQSPPLADELARAIVRIGDMPAAVAIGEHLVSEIYAVAV 120

RESULT 2
Q7TY70_MYCBO PRELIMINARY; PRT; 120 AA.
ID Q7TY70_MYCBO PRELIMINARY; PRT; 120 AA.
AC Q7TY70;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein Mb2661.
GN OrderedLocustNames=MB2661;
OS Mycobacterium bovis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium tuberculosis complex.
OX NCBI_TaxId=1765;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AF2122/97;
RX MEDLINE=22709107; PubMed=12788972; DOI=10.1073/pnas.1130426100;
RA Garrier T., Eglmeier K., Camus J.-C., Medina N., Manoor H.,
RA Pryor M., Dutfoy S., Grondin S., Lacroix C., Monsenpe C., Simon S.,
RA Harth B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA Parthill J., Barrell B.G., Cole S.T., Gordon S.V., Hewison R.G.;
RT "The complete genome sequence of Mycobacterium bovis."
RT Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882 (2003).
DR EMBL; BX248343; CAD94846.1; -; Genomic_DNA.
KW Complete proteome.
SQ SEQUENCE 120 AA; 13214 MW; 54287642A589EB28 CRC64;

Query Match 97.6%; Score 622; DB 2; Length 120;
Best Local Similarity 98.3%; Pred. No. 6.6e-54;
Matches 118; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MSTOPRRHSGIRAVGYAMAGRCGRIGRWGVHGOEAMNNLAIWHPKRVGSATTIYQVTDNRSH 60
Db 1 MSTOPRRHSGIRAVGYAMAGRCGRIGRWGVHGOEAMNNLAIWHPKRVGSATTIYQVTDNRSH 60
Qy 61 DQRTARVPDEITSTVSGWLSLGTQSPPLADELARAIVRIGDMPAAVAIGEHLVSEIYAVAV 120
Db 61 DQRTARVPDEITSTVSGWLSLGTQSPPLADELARAIVRIGDMPAAVAIGEHLVSEIYAVAV 120

RESULT 3
Q73WD2_MYCPA PRELIMINARY; PRT; 87 AA.
ID Q73WD2_MYCPA PRELIMINARY; PRT; 87 AA.
AC Q73WD2;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN OrderedLocustNames=MAP2728;
OS Mycobacterium paratuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium avium complex (MAC).
OX NCBI_TaxId=1770;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=k10;
RA Li L., Bannantine J., Zhang Q., Amonsin A., Alt D., Kapur V.;
RA Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB017237; AAS05045.1; -; Genomic_DNA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 87 AA; 9576 MW; 684CDBA0ED21034A CRC64;
```

```
Query Match 40.8%; Score 260; DB 2; Length 87;
Best Local Similarity 57.6%; Pred. No. 4.7e-18;
Matches 53; Conservative 12; Mismatches 17; Indels 10; Gaps 2;

Qy 28 RWGVHGOEAMNNLAIWHPKRVGSATTIYQVTDNRSHOQRTARVPDEITSTVSGWLSLGTQSP 87
Db 5 KMHVH-----N-PQRRPATIYQVTDNRSHOQRTARVPDEITSTVSGWLSLGTQSP 54

Qy 88 PLADELARAIVRIGDMPAAVAIGEHLVSEIYAVAV 119
Db 55 PLVDELERAVRGDMPAAVAIGEHLVSEIYAVAV 86

RESULT 4
Q5ILJ1_SILPO PRELIMINARY; PRT; 447 AA.
ID Q5ILJ1_SILPO PRELIMINARY; PRT; 447 AA.
AC Q5ILJ1_SILPO PRELIMINARY; PRT; 447 AA.
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DE 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE NADH:ubiquinone oxidoreductase, Na(+)-translocating, A subunit
DE (EC 1.6.5.-).
GN Name=nqra; OrderedLocustNames=SP0A0028;
OS Silicibacter pomeroyi.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales;
OC Rhodobacteraceae; Silicibacter.
OX NCBI_TaxId=89184;
RN [1]
RP NUCLEOTIDE SEQUENCE (LARGE SCALE GENOMIC DNA).
RC STRAIN=DS-3 / ATCC 700808 / DSM 15171;
RX PubMed=15602564; DOI=10.1038/nature03170;
RA Moran M.A., Buchan A., Gonzalez J.M., Heidelberg J.F., Whitman W.B.,
RA Kliene R.P., Henriksen J.R., King G.M., Belas R., Fugua C.,
RA Brinkac L.M., Lewis M., Johri S., Weaver B., Pai G., Eisen J.A.,
RA Rane E., Shelton W.M., Ye W., Miller T.R., Carlson J., Rasko D.A.,
RA Paulsen I.T., Ren Q., Daugherty S.C., Deboy R.T., Dodson R.J.,
RA Durkin A.S., Madupu R., Nelson W.C., Sullivan S.A., Rosovitz M.J.,
RA Haft D.H., Selengut J., Ward N.;
RT "Genome sequence of Silicibacter pomeroyi reveals adaptations to the
RT marine environment".
RL Nature 432:910-913 (2004).
DR EMBL; CP000032; AAV97168.1; -; Genomic_DNA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0016655; F:oxidoreductase activity; acting on NADH or . . . IEA.
DR GO; GO:0006814; P:sodium ion transport; IEA.
DR InterPro: IPR008703; NQRA.
DR Pfam: PF05896; NQRA; 1.
DR TIGFams; TIGR01936; nqra; 1.
KW Complete proteome; Oxidoreductase; Plasmid; Ubiquinone.
SQ SEQUENCE 447 AA; 47132 MW; BD7AD132AAB25EF6 CRC64;

Query Match 13.7%; Score 87.5; DB 2; Length 447;
Best Local Similarity 30.1%; Pred. No. 4;
Matches 34; Conservative 15; Mismatches 43; Indels 21; Gaps 5;

Qy 9 SGIRAVGYAMAGRCGRIGRWGVHGOEAMNNLAIW---HPRK--YQ-----SATTIYQV 56
Db 307 SGRQAGGFAYLGRARQVTL-IEEDRKQHTLGLPSPFAVQPVLSAFSKRLVLT 365

Qy 57 DRSHQRTARVPDEITSTVSGWLSLGTQSPPLADELARAIVRIGDMPAAVAIGEHLVSEIYAV 109
Db 366 SNLNGRGRAMP-----TGTFEELMPQVYLPQLLRALLVMDTDAQALG 410

RESULT 5
Q9K0B2_NEIMB PRELIMINARY; PRT; 691 AA.
ID Q9K0B2_NEIMB PRELIMINARY; PRT; 691 AA.
AC Q9K0B2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 24, Last annotation update)
```

```

DE Competence protein ComA.
GN Name=ComA; OrderedLocustNames=NM00702;
OS Neisseria meningitidis (serogroup B).
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxId=491;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MC58 / Serogroup B;
MEDLINE=20175755; PubMed=10710307; DOI=10.1126/science.287.5459.1809;
RA Tettein H., Saunders N.J., Heidelberg J.P., Jeffries A.C.,
RA Nelson K.E., Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F.,
RA Dodson R.J., Nelson W.C., Gwinn M.L., Deboy R.T., Peterson J.D.,
RA Hickey E.K., Haft D.H., Salzberg S.L., White O., Fleischmann R.D.,
RA Dougherty B.A., Mason T.M., Ciecko A., Parksey D.S., Blair E.,
RA Citron H., Clark E.B., Cotton M.D., Utterback T.R., Khouri H.M.,
RA Qin H., Vamathevan J.J., Gill J., Scarlato V., Maignani V., Pizzo M.,
RA Grandi G., Sun L., Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R.,
RA Venter J.C.;
RT "Complete genome sequence of Neisseria meningitidis serogroup B strain
MC58."
RL Science 287:1809-1815 (2000).
DR EMBL; AE002098; AAF4119.1; -; Genomic_DNA.
DR PIR; G8167; G8167.
DR TIGR; NM00702; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0030420; P:establishment of competence for transformation; IEA.
DR InterPro; IPR001279; Blactamase-like.
DR InterPro; IPR004477; ComEC_N-term.
DR InterPro; IPR004797; ComEC_Rec2.
DR Pfam; PF03772; Competence; 1.
DR Pfam; PF00753; Lactamase_B; 1.
DR TIGRFAMs; TIGR00360; ComEC_N-term; 1.
DR TIGRFAMs; TIGR00361; ComEC_Rec2; 1.
KW Complete proteome.
SQ
SEQUENCE 691 AA; 74408 MW; 5EC00CE0B166C6FE CRC64;

Query Match 13.4%; Score 85.5; DB 2; Length 691;
Best Local Similarity 27.1%; Pred. No. 10;
Matches 29; Conservative 14; Mismatches 51; Indels 13; Gaps 3;

QY 6 PHSIGIRAVGYPYAMAGRCRIGRWGQHQAAMNLAIMPRKY-----QSATIQ 54
DB 241 PQRSLVLMALAPYAMWRGRSLSAMATWQALAVLFDPSAVLVGVGTWLSFGLVALIWA 300
DB 301 CSGRLHEGRQYALRGQMAASVLSVLGLYLPASLPISPLVNAVAI 347

QY 55 VTRSDHSG-RTARVPGEITSTVS-GWLSLGTSQSLADELARAVRI 99
DB 301 CSGRLHEGRQYALRGQMAASVLSVLGLYLPASLPISPLVNAVAI 347

RESULT 6
Q9JVB8_NEIMA PRELIMINARY; PRT; 691 AA.
ID Q9JVB8_NEIMA PRELIMINARY; PRT; 691 AA.
AC Q9JVB8;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Competence protein.
GN Name=ComA; OrderedLocustNames=NMA0906;
OS Neisseria meningitidis (serogroup A).
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxId=56569;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=42491 / Serogroup A / Serotype 4A;
MEDLINE=2022556; PubMed=10761919; DOI=10.1038/3500665;
RA Parthill J., Achtman M., James K.D., Bentley S.D., Churcher C.M.,
RA Klee S.R., Morelli P., Basham D., Brown D., Chillingworth T.,
RA Davies R.M., Davis P., Devlin K., Felwell T., Hamlin N., Holroyd S.,
RA Jagsis K., Leacher S., Moule S., Mungall K.L., Quail M.A.,
RA Rajendram M.A., Rutherford K.M., Simmonds M., Skelton J.,
RA Whitehead S., Spratt B.G., Barrall B.G.;

```

```

RT "Complete DNA sequence of a serogroup A strain of Neisseria
meningitidis 22491."
RL Nature 404:502-506 (2000).
DR EMBL; AL162754; CAB84183.1; -; Genomic_DNA.
DR PIR; B81937; B81937.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0030420; P:establishment of competence for transformation; IEA.
DR GO; GO:006810; P:transport; IEA.
DR InterPro; IPR001279; Blactamase-like.
DR InterPro; IPR004477; ComEC_N-term.
DR InterPro; IPR004797; ComEC_Rec2.
DR Pfam; PF03772; Competence; 1.
DR Pfam; PF00753; Lactamase_B; 1.
DR TIGRFAMs; TIGR00360; ComEC_N-term; 1.
DR TIGRFAMs; TIGR00361; ComEC_Rec2; 1.
KW Complete proteome.
SQ
SEQUENCE 691 AA; 74435 MW; D4924B8F446D8D9 CRC64;

Query Match 13.1%; Score 83.5; DB 2; Length 691;
Best Local Similarity 26.2%; Pred. No. 16;
Matches 28; Conservative 15; Mismatches 51; Indels 13; Gaps 3;

QY 6 PHSIGIRAVGYPYAMAGRCRIGRWGQHQAAMNLAIMPRKY-----QSATIQ 54
DB 241 PQRSLVLMALAPYAMWRGRSLSAMATWQALAVLFDPSAVLVGVGTWLSFGLVALIWA 300
DB 301 CSGRLHEGRQYALRGQMAASVLSVLGLYLPASLPISPLVNAVAI 347

QY 55 VTRSDHSG-RTARVPGEITSTVS-GWLSLGTSQSLADELARAVRI 99
DB 301 CSGRLHEGRQYALRGQMAASVLSVLGLYLPASLPISPLVNAVAI 347

RESULT 7
Q5YS68_NOCFA PRELIMINARY; PRT; 331 AA.
ID Q5YS68_NOCFA PRELIMINARY; PRT; 331 AA.
AC Q5YS68;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Putative dehydrogenase.
GN OrderedLocustNames=nfa41240;
OS Nocardia farcinica.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Nocardiaceae; Nocardia.
OX NCBI_TaxId=37329;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=IFM 10152;
RX PubMed=15466710; DOI=10.1073/pnas.0406410101;
RA Ishikawa J., Yamashita A., Mikami Y., Hoshino Y., Kurita H., Hotta K.,
RA Shida T., Hattori M.;
RT "The complete genomic sequence of Nocardia farcinica IFM 10152."
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
DE EMBL; AP006618; BAD58973.1; -; Genomic_DNA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR InterPro; IPR002085; ADH_SF_Zn.
DR Pfam; PF00107; ADH_2inc_N; 1.
KW Complete proteome.
SQ
SEQUENCE 331 AA; 33227 MW; 8FE58428946EB29C CRC64;

Query Match 13.0%; Score 82.5; DB 2; Length 331;
Best Local Similarity 29.9%; Pred. No. 9;
Matches 38; Conservative 13; Mismatches 45; Indels 31; Gaps 6;

QY 10 GIRAVGYPYAMAGRCRIGRWGQHQAAMNLAIMPRKYQSATIQY-----TDRS---H 60
DB 84 GARVVGPPYAGAGIGGGVPTGGAERAL-----VTBOTLLAVPAGLDDRAVALVH 133
QY 61 DGRTRAV-----PSGEIST--VSGWLSLGTSQ--SPLADELARAVRIGDPAAYAI 108
DB 134 DGRITALAPDRATVQPGDLVLTVAAGGLGTLTOLAHRAQAKVVAARGAKELEAHL 193

```



```

RESULT 11
O7CVT0_AGR75      PRT;      520 AA.
ID O7CVT0_AGR75
AC O7CVT0;
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DE 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE AGR_L_242P.
GN OrderedLocuNames=AGR_L_242;
OC Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiales; Rhizobium/Agrobacterium group; Agrobacterium.
OX NCBI_TaxID=176299;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=21608551; PubMed=11743194; DOI=10.1126/science.1066803;
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M., Mullin L.,
RA Qurollo B., Goldman B.S., Cao Y., Akenazi M., Halling C., Mullin L.,
RA Hounmel K., Gordon J., Vaudin M., Iatchouk O., Epp A., Liu F.,
RA Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,
RA Flanagan C., Crowell C., Guron J., Lomo C., Sear C., Strub G.,
RA Cielo C., Slater S.;
RT "Genome sequence of the plant pathogen and biotechnology agent
RT Agrobacterium tumefaciens C58."
RL Science 294:2323-2328(2001).
CC -1- SIMILARITY: Belongs to the class-III pyridoxal-phosphate-dependent
CC      aminotransferase family.
CC      EMBL: AE008209; AAK8689.1; -; Genomic DNA.
DR GO: GO:0030170; P:pyridoxal phosphate binding; IEA.
DR GO: GO:0008483; F:transaminase activity; IEA.
DR InterPro: IPR005814; AminoTrans_3.
DR PANTHER: PTHR11986; AminoTrans_3.
DR Pfam: PF00202; AminoTrans_3; 1.
DR PROSITE: PS00600; AA_TRANSER_CLASS_3; 1.
KW Pyridoxal phosphate.
SQ SEQUENCE      520 AA; 56284 MW; BE51962738F5FDE CRC64;

Query Match      12.6%; Score 80.5; DB 2; Length 520;
Best Local Similarity 28.0%; Pred. No. 23;
Matches 28; Conservative 9; Mismatches 56; Indels 7; Gaps 2;

QY      2 STORPHSGIRAVPYAMAGRCGRIGRWGVHQAAMNLAIWHRKYQASATTIVTDRSHD 61
DB      8 SPBRPFHSA--AGNNQALAAARAVMASWINSNPMGASAIWPERISPMTIDIKIAEMD 65
QY      62 GRTARVPGEITSTVSGWLSLCTQSPLDELARARIGD 101
DB      66 RNAVHPFTQLKDFASGKLGE-----PTIVETGKGIRIOD 100

RESULT 12
COMA_NEIGO
ID COMA_NEIGO      STANDARD;      PRT;      691 AA.
AC P51973;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Competence protein coma.
GN Name=coma;
OS Neisseria gonorrhoeae.
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=485;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=95020543; PubMed=7934834;
RA Facius D., Meyer T.F.;
RT "A novel determinant (coma) essential for natural transformation
RT competence in Neisseria gonorrhoeae and the effect of a coma defect on
RT plin variation."
RT Mol. Microbiol. 10:699-712(1993).
RL [2]
RP NUCLEOTIDE SEQUENCE.

```

```

RC STRAIN=MS11A;
RX MEDLINE=96249702; PubMed=8830266;
RA Fueseneger M., Facius D., Meier J., Meyer T.F.;
RT "A novel peptidoglycan-linked lipoprotein (ComL) that functions in
RT natural transformation competence of Neisseria gonorrhoeae."
RL Mol. Microbiol. 19:1095-1105(1996).
CC -1- FUNCTION: Essential for natural transformation. Could be a
CC      transporter involved in DNA uptake.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC      (Probable).
CC -1- SIMILARITY: To B.subtilis comEC, H.influenzae REC2, and E.coli
CC      ycal.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL: S75490; AAB32261.1; -; Genomic DNA.
DR EMBL: Z49895; -; NOT_ANNOTATED_CDS; Genomic DNA.
DR PIR: S39867; S39867
DR InterPro: IPR001279; Blactamase-like.
DR InterPro: IPR004477; ComEC_N-term.
DR InterPro: IPR004797; ComEC_Rec2.
DR Pfam: PF03772; Competence; 1.
DR Pfam: PF00753; Lactamase B; 1.
DR TIGRPFAM: TIGR00360; ComEC_N-term; 1.
DR TIGRPFAM: TIGR00361; ComEC_Rec2; 1.
KW Competence; Inner membrane; Membrane; Transmembrane; Transport.
FT TRANSMEM      183      203      Potential.
FT TRANSMEM      220      240      Potential.
FT TRANSMEM      280      300      Potential.
FT TRANSMEM      322      342      Potential.
FT TRANSMEM      347      367      Potential.
FT TRANSMEM      396      416      Potential.
SQ SEQUENCE      691 AA; 74358 MW; CBB8CC7601A347FE CRC64;

Query Match      12.5%; Score 79.5; DB 1; Length 691;
Best Local Similarity 26.2%; Pred. No. 40;
Matches 28; Conservative 14; Mismatches 52; Indels 13; Gaps 3;

QY      6 PRHSGIRAVPYAMAGRCGRIGRWGVHQAAMNLAIWHRKY-----QSATTIVQ 54
DB      241 PQGSAVLMAAPAMAMRGRLSAMATWQALAAVLLFDPLAVLGVTWLSFGLVALIWA 300
QY      55 YVDRSHDG-RTARVPGEITSTVS-GWLSLCTQSPLDELARARIGD 99
DB      301 CAGRLVEGKRGRTAVRGQMAASVLSVLGLYLPASLPLVSPLVNAVSI 347

RESULT 13
O5KXV7_GEOKA
ID O5KXV7_GEOKA      PRELIMINARY;      PRT;      270 AA.
AC O5KXV7;
DT 01-FEB-2005 (TREMBLrel. 29, Created)
DT 01-FEB-2005 (TREMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TREMBLrel. 29, Last annotation update)
DE Hypothetical conserved protein.
GN OrderedLocuNames=GK2874;
OC Geobacillus kaustophilus.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Geobacillus.
OX NCBI_TaxID=1462;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=HTA426;
RX PubMed=15576355; DOI=10.1093/nar/gkh970;
RA Takami H., Takaki Y., Chee G.-J., Nishi S., Shimamura S., Suzuki H.,
RA Matsui S., Uchiyama I.;
RT "Thermoadaptation trait revealed by the genome sequence of
RT thermophilic Geobacillus kaustophilus."
RT Nucleic Acids Res. 32:6292-6303(2004).
CC -1- CATALYTIC ACTIVITY: Release of N-terminal proline from a peptide.

```

DR EMBL; BA000043; BAD77159.1; -; Genomic_DNA.
 DR GO; GO:0004177; F:aminopeptidase activity; IEA.
 DR GO; GO:0016787; F:hydrolase activity; IEA.
 DR GO; GO:0006725; P:aromatic compound metabolism; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR000073; A/b hydrolase.
 DR InterPro; IPR003089; AB_hydrolase.
 DR InterPro; IPR000639; Epox_hydrolase.
 DR InterPro; IPR008262; Lipase_AS.
 DR InterPro; IPR002410; Peptidase_S33.
 DR InterPro; IPR000379; Ser esters_
 DR Pfam; PF00561; Abhydrolase_1; 1.
 DR PRINTS; PRO0111; ABHYDROLASE.
 DR PRINTS; PRO0412; EPOXYHYDLASE.
 DR PRINTS; PRO0793; PROAMNOPTASE.
 DR PROSITE; PS00120; LIPASE_SER; UNKNOWN_1.
 KW Aminopeptidase; Complete proteome; Hydrolase; Hypothetical protein.
 SQ SEQUENCE 270 AA; 30278 MW; 9475BD73C480910 CRC64;

Query Match 12.4%; Score 79; DB 2; Length 270;
 Best Local Similarity 23.6%; Pred. No. 16;
 Matches 25; Conservative 16; Mismatches 41; Indels 24; Gaps 4;

QY 17 YMAAGRCGRIGRWGHOEAMM-----LATMHPKQVSATITYQTD--RSHDGR 64
 DB 70 HAAADLALLDDEMGVEQVAVVLCYSGMGRLLAFVAVMHPRRVRLVLESSPGLTEEEER 129
 QY 65 AVVPGE-----ITSTVSGW--LSELGTQSPLELARA 98
 DB 130 ARREADEALARKIETEGVAAPFDDMKIPLFATQQLPFPVAAR 175

RESULT 14
 Q3MFP2_AMYMD PRELIMINARY; PRT; 163 AA.
 ID Q3MFP2_AMYMD PRELIMINARY; PRT; 163 AA.
 AC Q3MFP2;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein (Fragment).
 OS Amycolatopsis mediterranei (Nocardia mediterranei).
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Pseudonocardiales; Pseudonocardaceae; Amycolatopsiis.
 OX NCBI_TaxId=33910;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=U32;
 RA Wang W.W., Gao J., Hiao J., Zhao G., Jiang W.;
 RA "Novel Two-Component System amrB-amkB Involved in the Regulation of
 RT Central Carbohydrate Metabolism in Rifamycin SV-Producing
 RT Amycolatopsis mediterranei U32.";
 RL Curr. Microbiol. 48:14-19(2004).
 DR EMBL; AJ318385; CAC42482.1; -; Genomic_DNA.
 KW Hypothetical protein.
 FT NON TER 1
 SQ SEQUENCE 163 AA; 18071 MW; 1546D37007E76ED1 CRC64;

Query Match 12.2%; Score 78; DB 2; Length 163;
 Best Local Similarity 29.1%; Pred. No. 12;
 Matches 25; Conservative 7; Mismatches 40; Indels 14; Gaps 2;

QY 6 PHSGRFVAVPYAMAGRCGRIGRWGHOEAMMMLAIHPKQVSATITYQTD RSHDGR 65
 DB 4 PRGTG--ATRPVPRARHRRKLTGRKXDEANSLAMKVMHXSRYQ-----DDHYA 49
 QY 66 RVPGDEITSTVSGWSELGTQSPLEAD 91
 DB 50 DVAGTEQVWVTGQDWLSRFGLPPAG 75

RESULT 15
 Q88E24_PSEPK

ID Q88E24_PSEPK PRELIMINARY; PRT; 456 AA.
 AC Q88E24;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE DNA repair protein Rada.
 GN Name=rdh; OrderedLocustNames=PP4644;
 OS Pseudomonas putida (strain KT2440).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 OX NCBI_TaxId=160488;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=22423060; PubMed=12534463;
 RX DOI=10.1046/j.1462-2920.2002.00366.x;
 RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,
 RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,
 RA Brinkac L.M., Beanan M.J., DeBoy R.T., Daugherty S.C., Kolonay J.F.,
 RA Madupu R., Nelson W.C., White O., Peterson J.D., Khoult H.M.,
 RA Hance I., Chris Lee P., Holtzapfle E.K., Scanlan D., Tran K.,
 RA Moazzez A., Uetzerback T.R., Rizzo M., Lee K., Kosack D., Moestl D.,
 RA Medler H., Lauber J., Stjepandic D., Hohnesiel J., Straetz M., Helm S.,
 RA Kiewitz C., Eisen J.A., Timmis K.N., Duesterhoeft A., Tuemmler B.,
 RA Fraser C.M.;
 RA "Complete genome sequence and comparative analysis of the
 RT metabolically versatile Pseudomonas putida KT2440.";
 RT Environ. Microbiol. 4:799-808(2002).
 DR EMBL; AE016791; AAN70217.1; -; Genomic_DNA.
 DR TIGR; PP4644; -.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0004176; F:ATP-dependent peptidase activity; IEA.
 DR GO; GO:0003684; F:damaged DNA binding; IEA.
 DR GO; GO:0008094; F:DNA-dependent ATPase activity; IEA.
 DR GO; GO:0017111; F:nucleoside-triphosphatase activity; IEA.
 DR GO; GO:0000166; F:nucleotide binding; IEA.
 DR GO; GO:0004252; F:serine-type endopeptidase activity; IEA.
 DR GO; GO:0006310; P:DNA recombination; IEA.
 DR GO; GO:0006281; P:DNA repair; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR003593; AAA_ATPase.
 DR InterPro; IPR001984; Peptidase_S16.
 DR InterPro; IPR004504; Rada.
 DR InterPro; IPR001553; Rech.
 DR PRINTS; PRO0830; ENDOLAPTASE.
 DR SMART; SM00382; AAA; 1.
 DR TIGRFAMs; TIGR00416; smg; 1.
 DR PROSITE; PS0162; RECA_2; 1.
 KW Complete proteome.
 SQ SEQUENCE 456 AA; 48741 MW; 36CF6E54B5FSD265 CRC64;

Query Match 12.2%; Score 77.5; DB 2; Length 456;
 Best Local Similarity 23.3%; Pred. No. 40;
 Matches 30; Conservative 15; Mismatches 43; Indels 41; Gaps 6;

QY 19 WAGRCGRIGRWGHOEAMM-----NLAIHPKQVSATITYQTD----- 56
 DB 20 WAGQCGEGGANVTLETETIESGGAAPSSGAGWGGQAQIKTLAEVSEIEI PRETTIS 79
 QY 57 --DR-----SHDGRFVAVPYAMAGRCGRIGRWGSELGTQSPLELARA 109
 DB 80 ELDRVLGGGLVDGVSIVLIGD-----PGIKSTIILQTLCLN-IAVG-MPALVYTG 127
 QY 110 EHLSEVAV 118
 DB 128 EESQGVAM 136

Search completed: March 23, 2006, 05:24:32
 Job time : 31.6915 secs

NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 25258
LENGTH: 510
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25258

Query Match 11.8%; Score 75; DB 2; Length 510;
Best Local Similarity 26.3%; Pred. No. 2.6;
Matches 31; Conservative 19; Mismatches 52; Indels 16; Gaps 4;

QY 11 TAAVGYAAGRCGRGNGVHOEAMNLAIMHPRVGSATITVQVDRSHDRTARV--- 67
DB 261 IEAAVPEAQSGLP-----QVGTQAVRLAAPEGEVLQARIVSILPEANRDRTRVARVLE 314
QY 68 ---PGDEITSTVSGMLSELTGSP-LADELARAVRIGDMPAAVYAI---GEHLSEVETAV 118
DB 315 MANPGORLRKMGSGQIALKGNQEPALLVPESEAVIRIKKALAYVVDGPKFHPVEVQV 372

RESULT 3

US-09-252-991A-26360
Sequence 26360, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
PRIOR FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 26360
LENGTH: 330
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26360

Query Match 11.6%; Score 74; DB 2; Length 330;
Best Local Similarity 25.2%; Pred. No. 1.9;
Matches 31; Conservative 13; Mismatches 59; Indels 20; Gaps 4;

QY 2 STQGRHSGIRAVGPYAA-----GRCGRIGRW-----GVHGEAMNLAIMH 43
DB 26 TSTRRPGEPLPMAAPVAMARRRLGSRSSRSRGWPSRVCTTMTPTPRAOSSAMIAIGA 85
QY 44 PKVOSATITVQVDRS-HDGRTRAVPGDEITSTVSGMLSELTGSP-LADELARAVRIGD 102
DB 66 RRRRLTSIPMSAVPPEAQSRCMSITSSAVCSRSSCVRLTAQPTCTPR-RRCIRRCRG 144
QY 103 PAA 105
DB 145 PAA 147

RESULT 4

US-10-104-047-2749
Sequence 2749, Application US/10104047
Patent No. 6943241
GENERAL INFORMATION:
APPLICANT: HELIX RESEARCH INSTITUTE
TITLE OF INVENTION: NO. 6943241el full length CDNA
FILE REFERENCE: H1-A0105
CURRENT APPLICATION NUMBER: US/10/104,047
PRIOR FILING DATE: 2002-03-25
PRIOR APPLICATION NUMBER:
NUMBER OF SEQ ID NOS: 4096
SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 2749
LENGTH: 449
TYPE: PRT
ORGANISM: Homo sapiens
US-10-104-047-2749

Query Match 11.5%; Score 73; DB 2; Length 449;
Best Local Similarity 28.1%; Pred. No. 3.8;
Matches 18; Conservative 15; Mismatches 27; Indels 4; Gaps 1;

QY 54 QVDRSHDGRTRAVPGDEITSTVSGMLSELTGSP-LADELARAVRIGDMPAAVYAI 113
DB 249 ELSDVSGGSKATTP-----ASTANSDVATITPTDPLKEENGFFVKTDAFNKSEISKIIE 304
QY 114 VEIA 117
DB 305 VQVA 308

RESULT 5

US-09-949-016-9731
Sequence 9731, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
PRIOR FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 9731
LENGTH: 868
TYPE: PRT
ORGANISM: Human
US-09-949-016-9731

Query Match 11.5%; Score 73; DB 2; Length 868;
Best Local Similarity 28.1%; Pred. No. 9.1;
Matches 18; Conservative 15; Mismatches 27; Indels 4; Gaps 1;

QY 54 QVDRSHDGRTRAVPGDEITSTVSGMLSELTGSP-LADELARAVRIGDMPAAVYAI 113
DB 186 ELSDVSGGSKATTP-----ASTANSDVATITPTDPLKEENGFFVKTDAFNKSEISKIIE 241
QY 114 VEIA 117
DB 242 VQVA 245

RESULT 6

US-08-737-526-4
Sequence 4, Application US/08737526
Patent No. 5871966
GENERAL INFORMATION:
APPLICANT: Kofoed, Lene Venke
APPLICANT: Andersen, Lene No. 5871966boe
APPLICANT: Kauppinen, Markus Sakari
APPLICANT: Christgau, Stephan
TITLE OF INVENTION: An Enzyme with Endo-1,3(4)-B-Glucanase
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESS: No. 5871966o No. 5871966disk of No. 5871966th America, Inc.
STREET: 405 Lexington Avenue
CITY: New York

STATE: NY
COUNTRY: USA
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/737,526
FILING DATE: 08-NOV-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Valecia, Gregg A
REGISTRATION NUMBER: 35,127
REFERENCE/DOCKET NUMBER: 4174.204-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 292 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-737-526-4

Query Match 11.1%; Score 71; DB 1; Length 292;
Best Local Similarity 26.5%; Pred. No. 3.7;
Matches 35; Conservative 18; Mismatches 51; Indels 28; Gaps 8;

QY 4 QPRHSGIRAVGPYVMAGRG---RIGRWGVHQAEMNLAIWPKVKQSATTIYQVTDSSH 60
21 QPPTSGFNLVWTDTPFAGNGGTSPNQNNNI---ITGNLVNAEQETYSSTANV--QLS 75

DB 61 DGRTR-VP---GDEITSTVSGM---LSELGTQSLADELAR---AVRIGD----- 101
76 GGSITQLVPMWDSKGTSTFGMTSGRLSKYTFPPAAGKVTRELAIRFGSNAQNAKOG 135

QY 102 -WPAVVAIGEHL 112
DB 136 IWPFAWMLGDSL 147

RESULT 7
US-09-098-580-4
Sequence 4, Application US/09098580
Patent No. 6140096
GENERAL INFORMATION:
APPLICANT: Kofod, Lene Venke
APPLICANT: Andersen, Lene No. 6140096boe
APPLICANT: Kauppinen, Markus Sakari
TITLE OF INVENTION: An Enzyme with Endo-1,3(4)-B-Glucanase
TITLE OF INVENTION: Activity
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 61400960 No. 6140096disk of No. 6140096th America, Inc.
STREET: 405 Lexington Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/098,580
FILING DATE:
CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/737,526
FILING DATE: 08-NOV-1996
ATTORNEY/AGENT INFORMATION:
NAME: Valecia, Gregg A
REGISTRATION NUMBER: 35,127
REFERENCE/DOCKET NUMBER: 4174.204-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 292 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-098-580-4

Query Match 11.1%; Score 71; DB 2; Length 292;
Best Local Similarity 26.5%; Pred. No. 3.7;
Matches 35; Conservative 18; Mismatches 51; Indels 28; Gaps 8;

QY 4 QPRHSGIRAVGPYVMAGRG---RIGRWGVHQAEMNLAIWPKVKQSATTIYQVTDSSH 60
21 QPPTSGFNLVWTDTPFAGNGGTSPNQNNNI---ITGNLVNAEQETYSSTANV--QLS 75

DB 61 DGRTR-VP---GDEITSTVSGM---LSELGTQSLADELAR---AVRIGD----- 101
76 GGSITQLVPMWDSKGTSTFGMTSGRLSKYTFPPAAGKVTRELAIRFGSNAQNAKOG 135

QY 102 -WPAVVAIGEHL 112
DB 136 IWPFAWMLGDSL 147

RESULT 8
US-08-801-263A-5
Sequence 5, Application US/08801263A
Patent No. 5811407
GENERAL INFORMATION:
APPLICANT: Johnston, Robert E.
APPLICANT: Davis, Nancy L.
APPLICANT: Simpson, Dennis A.
TITLE OF INVENTION: System for the In Vivo Delivery and
TITLE OF INVENTION: Expression of Heterologous Genes in the Bone Marrow
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Bell Seltzer Park & Gibson, P.A.
STREET: 1211 East Morehead Street
CITY: Charlotte
STATE: No. 5811407th Carolina
COUNTRY: USA
ZIP: 28234
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/801,263A
FILING DATE: 19-FEB-1997
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Sibley, Kenneth D.
REGISTRATION NUMBER: 31,665
REFERENCE/DOCKET NUMBER: 5470-147
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-420-2200
TELEFAX: 919-881-3175
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 2517 amino acids

TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-801-263A-5

Query Match 11.0%; Score 70; DB 1; Length 2517;

Best Local Similarity 20.3%; Pred. No. 87;
Matches 32; Conservative 19; Mismatches 47; Indels 60; Gaps 7;

QY 5 RPRHSGIRAVGPYAMAGRCGRIGRWGVHQAEMNNLAIW--HPRK-VQSATIIYQVTDK--- 58
DB 810 RPRHNVVLCGDP-----KQCGFFN-----MMQKVFNNHPEKDICTKTFYKFSRRT 857
QY 59 -----SHDGR-----TARVPGDEITSTVSGWLSLGTQSPLA 90
DB 858 QPVTAVSTLHYDGMKMTNPKCKNIEIDITGATPKPKEDIIITCFRGMVQLOIDYDGH 917
QY 91 DELARAV-----RIGDMPAAVAIGEHLISV 114
DB 918 EVMTAAASQGLTRKGVAVRQKVNENPLYAITSEHNV 955

RESULT 9

US-09-102-248-5
Sequence 5, Application US/09102248
Patent No. 6008035

GENERAL INFORMATION:

APPLICANT: Johnston, Robert E.

APPLICANT: Davis, Nancy L.

APPLICANT: Simpson, Dennis A.

TITLE OF INVENTION: System for the In Vivo Delivery and

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSEE: Bell Seltzer Park & Gibson, P.A.

STREET: 1211 East Morehead Street

CITY: Charlotte

STATE: No. 6008035th Carolina

COUNTRY: USA

ZIP: 28234

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/102,248

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/801,263

FILING DATE: 19-FEB-1997

ATTORNEY/AGENT INFORMATION:

NAME: Sibley, Kenneth D.

REGISTRATION NUMBER: 31,665

REFERENCE/DOCKET NUMBER: 5470-147

TELECOMMUNICATION INFORMATION:

TELEPHONE: 919-420-2200

TELEFAX: 919-881-3175

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 2517 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-09-102-248-5

Query Match 11.0%; Score 70; DB 2; Length 2517;

Best Local Similarity 20.3%; Pred. No. 87;
Matches 32; Conservative 19; Mismatches 47; Indels 60; Gaps 7;

QY 5 RPRHSGIRAVGPYAMAGRCGRIGRWGVHQAEMNNLAIW--HPRK-VQSATIIYQVTDK--- 58
DB 918 EVMTAAASQGLTRKGVAVRQKVNENPLYAITSEHNV 955

DB 810 RPRHNVVLCGDP-----KQCGFFN-----MMQKVFNNHPEKDICTKTFYKFSRRT 857

QY 59 -----SHDGR-----TARVPGDEITSTVSGWLSLGTQSPLA 90

DB 858 QPVTAVSTLHYDGMKMTNPKCKNIEIDITGATPKPKEDIIITCFRGMVQLOIDYDGH 917

QY 91 DELARAV-----RIGDMPAAVAIGEHLISV 114

DB 918 EVMTAAASQGLTRKGVAVRQKVNENPLYAITSEHNV 955

RESULT 10

US-09-367-764-5

Sequence 5, Application US/09367764

Patent No. 6583121

GENERAL INFORMATION:

APPLICANT: Johnston, Robert E.

APPLICANT: Davis, Nancy L.

APPLICANT: Simpson, Dennis A.

TITLE OF INVENTION: System for the In Vivo Delivery and

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSEE: Bell Seltzer Park & Gibson, P.A.

STREET: 1211 East Morehead Street

CITY: Charlotte

STATE: No. 6583121th Carolina

COUNTRY: USA

ZIP: 28234

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/367,764

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/801,263

FILING DATE: 19-FEB-1997

ATTORNEY/AGENT INFORMATION:

NAME: Sibley, Kenneth D.

REGISTRATION NUMBER: 31,665

REFERENCE/DOCKET NUMBER: 5470-147

TELECOMMUNICATION INFORMATION:

TELEPHONE: 919-420-2200

TELEFAX: 919-881-3175

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 2517 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-09-367-764-5

Query Match 11.0%; Score 70; DB 2; Length 2517;

Best Local Similarity 20.3%; Pred. No. 87;
Matches 32; Conservative 19; Mismatches 47; Indels 60; Gaps 7;

QY 5 RPRHSGIRAVGPYAMAGRCGRIGRWGVHQAEMNNLAIW--HPRK-VQSATIIYQVTDK--- 58
DB 810 RPRHNVVLCGDP-----KQCGFFN-----MMQKVFNNHPEKDICTKTFYKFSRRT 857
QY 59 -----SHDGR-----TARVPGDEITSTVSGWLSLGTQSPLA 90
DB 858 QPVTAVSTLHYDGMKMTNPKCKNIEIDITGATPKPKEDIIITCFRGMVQLOIDYDGH 917
QY 91 DELARAV-----RIGDMPAAVAIGEHLISV 114
DB 918 EVMTAAASQGLTRKGVAVRQKVNENPLYAITSEHNV 955

RESULT 11
US-09-252-991A-27332
; Sequence 27332, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27332
; LENGTH: 329
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27332

Query Match 10.9%; Score 69.5; DB 2; Length 329;
Best Local Similarity 30.4%; Pred. No. 6.5; Mismatches 55; Indels 13; Gaps 5;
Matches 35; Conservative 12;

QY 4 ORPHSGIRAVGPFYAMAGRCGRIGRWGHOEAMMNLAIW-HPRKVQSATIIYQVTDNRSHD 61
DB 216 QRPAGRG-RALVPRARG-FARLRLSGAGQGRPARVWRGRRRRCRAGACVSRARG 273
QY 62 GRT-----ARVPGDEITSTVSGWSELGTOSPLADELARAVRIGDWPAAVAIGE 110
DB 274 QRTDGRARCRPRADRLSSAVGALS---GEPALAGQAPVAGGQGPQAGHE 325

RESULT 12
US-09-252-991A-31264
; Sequence 31264, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 31264
; LENGTH: 630
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31264

Query Match 10.9%; Score 69.5; DB 2; Length 630;
Best Local Similarity 28.2%; Pred. No. 15; Mismatches 35; Indels 7; Gaps 3;
Matches 22; Conservative 14;

QY 43 HPRKVQSATIIYQVTDNRSH--DGRTRAVPGDEITSTVSGWSELGTOSPLADELARAVRI 99
DB 165 HWPSSSATELSGARPARHAQSKTLPIDPKLAETRAVSRP-----ERADETLAAENV 240
QY 100 GDMPAVAIAGEHLSVEIA 117
DB 241 SDMAQSYSGTYNAKLA 258

RESULT 13
US-09-489-039A-9779
; Sequence 9779, Application US/09489039A

; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 9779
; LENGTH: 625
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-9779

Query Match 10.8%; Score 69; DB 2; Length 625;
Best Local Similarity 22.9%; Pred. No. 18; Mismatches 33; Indels 30; Gaps 3;
Matches 22; Conservative 11;

QY 28 RWGVHQAAMNL-----ALMHPKVQSATIIYQVTDNRSHDGRTRAVPGDEITSTV 76
DB 443 QMGKAVSLMEIPTGTETLIDNIVCFWPEKAVKA-----GDEIDFRY 484
QY 77 S-GWSELGTOSPLADELARAVRIGDWPAAVAIGE 111
DB 485 RLYWSAQPVSTPLARVLATRTGCGFPEGMAFGSH 520

RESULT 14
US-09-477-962-118
; Sequence 118, Application US/09477962
; Patent No. 6927286
; GENERAL INFORMATION:
; APPLICANT: SHEN, BEN
; APPLICANT: DU, LIANGCHENG
; APPLICANT: SANCHEZ, CESAR
; APPLICANT: CHEN, MEI
; APPLICANT: EDWARDS, DANIEL J.
; TITLE OF INVENTION: BLEOMYCIN GENE CLUSTER COMPONENTS AND THEIR USES
; FILE REFERENCE: 407T-895820US
; CURRENT APPLICATION NUMBER: US/09/477,962
; PRIOR FILING DATE: 2000-01-05
; PRIOR APPLICATION NUMBER: 60/115,435
; PRIOR FILING DATE: 1999-01-06
; PRIOR APPLICATION NUMBER: 60/118,848
; PRIOR FILING DATE: 1999-02-05
; NUMBER OF SEQ ID NOS: 133
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 118
; LENGTH: 451
; TYPE: PRT
; ORGANISM: Streptomyces verticillius
; FEATURE:
; OTHER INFORMATION: ORF33
US-09-477-962-118

Query Match 10.8%; Score 68.5; DB 2; Length 451;
Best Local Similarity 27.8%; Pred. No. 13; Mismatches 46; Indels 19; Gaps 4;
Matches 30; Conservative 13;

QY 6 PHHSGIRAVGPFYAMAGRCGRIGRWGHOEAMMNLAIWHPKVQSATIIYQVTDNRSHDGRTA 65
DB 58 PHHAG-----LGRVHSSQAALVAL-GPMVASATITGVAAGDGHQAGRRG 101
QY 66 RVPGEITSTVSGWSELGTOS-PLADELARAVRIGDWPAAVAIGEHL 112
DB 102 QVTGVGVRVAVAGGDELDGAE DRPHARG--RLDDGLWMAAERLADLL 147

RESULT 15
US-09-976-594-699

```
; Sequence 699, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
; APPLICANT: Furness, Michael
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/976,594
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 699
; LENGTH: 544
; TYPE: PR1
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Inocyte ID No. 6673549 4113161CD1
US-09-976-594-699
```

```
Query Match 10.8%; Score 68.5; DB 2; Length 544;
Best Local Similarity 24.8%; Pred. No. 17;
Matches 37; Conservative 16; Mismatches 43; Indels 53; Gaps 9;

QY 10 GIRAVGPYVWAGRCGRIGRW-----GVHDEAMNNLAIWHPKVVQSATIVQVTD 58
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 25 GALGPGPHRW-GPCG--GSMWQKFKYQDGPGRGLGSEDIRRAREARPKTPRP---QLSDR 78
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 59 SHDGTARVPDEITSTVS-----GWLSELTQ-----SPLAD 91
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 79 S---RRKKVPASRISRLANFGGLAVGLGLVLAEMAKSMPGRLQSEGSGLDSSPFLS 135
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 92 ELRAVRIGDW-----PAAVAIGEHLSVE 115
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 136 E-AMMERIVQTLCTVRGAAALKVGMLSIQ 163
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

Search completed: March 23, 2006, 05:30:20
Job time : 9.2844 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: March 23, 2006, 06:14:33 ; Search time 25.2352 Seconds
(without alignments)
1966.885 Million cell updates/sec

Title: US-10-617-038-30

Perfect score: 637
Sequence: 1 MSTOPRRHSGIRAVGPYMA.....DMPAAVAIGEHLSVEIAVAV 120Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 200000000Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_AA_Main:*
1: /cgn2_6/prodata/1/pubppaa/US07_PUBCOMB.pep:*
2: /cgn2_6/prodata/1/pubppaa/US08_PUBCOMB.pep:*
3: /cgn2_6/prodata/1/pubppaa/US09_PUBCOMB.pep:*
4: /cgn2_6/prodata/1/pubppaa/US10_PUBCOMB.pep:*
5: /cgn2_6/prodata/1/pubppaa/US10_PUBCOMB.pep:*
6: /cgn2_6/prodata/1/pubppaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	637	100.0	120	4 US-10-617-038-30	Sequence 30, Appl
2	80	12.6	551	4 US-10-425-114-39387	Sequence 39387, A
3	79.5	12.5	1167	5 US-10-794-514A-328	Sequence 328, App
4	77.5	12.2	456	4 US-10-282-122A-68099	Sequence 68099, A
5	77	12.1	278	4 US-10-424-599-217918	Sequence 217918,
6	75.5	11.9	272	4 US-10-437-963-170212	Sequence 170212,
7	75	11.8	8605	6 US-11-097-143-28128	Sequence 28128, A
8	74.5	11.7	279	4 US-10-437-963-149941	Sequence 149941,
9	74	11.6	284	4 US-10-156-761-8113	Sequence 8113, Ap
10	74	11.6	535	4 US-10-282-122A-62103	Sequence 62103, A
11	73	11.5	69	3 US-09-864-761-43789	Sequence 43789, A
12	73	11.5	258	4 US-10-106-698-6371	Sequence 6371, Ap
13	73	11.5	365	4 US-10-732-923-10308	Sequence 10308, A
14	73	11.5	449	4 US-10-104-047-2749	Sequence 2749, Ap
15	73	11.5	715	3 US-09-815-242-10450	Sequence 10450, A
16	73	11.5	715	4 US-10-369-493-23660	Sequence 23660, A
17	73	11.5	715	4 US-10-282-122A-43223	Sequence 43223, A
18	73	11.5	715	4 US-10-282-122A-75843	Sequence 75843, A
19	73	11.5	766	4 US-10-205-219-157	Sequence 157, App
20	73	11.5	1277	4 US-10-276-774-2154	Sequence 2154, Ap
21	72.5	11.4	210	4 US-10-437-963-147829	Sequence 147829,
22	72.5	11.4	492	5 US-10-450-763-45094	Sequence 45094, A
23	72.5	11.4	472	5 US-10-156-761-10210	Sequence 10210, A
24	71	11.1	472	5 US-10-732-923-1072	Sequence 1072, Ap
25	71	11.1	19723	4 US-10-084-846A-5	Sequence 5, Appl
26	70.5	11.1	99	4 US-10-437-963-159656	Sequence 159656,
27	70.5	11.1	281	4 US-10-369-493-4003	Sequence 4003, Ap

28	70.5	11.1	753	4 US-10-282-122A-49874	Sequence 49874, A
29	70.5	11.1	954	4 US-10-437-963-184007	Sequence 184007,
30	70	11.0	838	5 US-10-779-597-95	Sequence 95, Appl
31	70	11.0	3362	4 US-10-378-083-6	Sequence 6, Appl
32	69.5	10.9	235	5 US-10-481-265-19	Sequence 19, Appl
33	69.5	10.9	330	4 US-10-424-599-153232	Sequence 153232,
34	69.5	10.9	453	4 US-10-156-761-1136	Sequence 14336, A
35	69.5	10.9	696	3 US-09-738-628-5259	Sequence 5259, Ap
36	69.5	10.9	981	4 US-10-156-761-15071	Sequence 15071, A
37	69	10.8	327	4 US-10-437-963-201707	Sequence 201707,
38	69	10.8	391	4 US-10-156-761-9791	Sequence 9791, Ap
39	69	10.8	476	4 US-10-425-114-52031	Sequence 52031, A
40	69	10.8	646	4 US-10-437-963-136807	Sequence 136807,
41	69	10.8	921	4 US-10-467-685-14	Sequence 14, Appl
42	69	10.8	927	4 US-10-408-765A-2938	Sequence 2938, Ap
43	69	10.8	1380	5 US-10-450-763-33073	Sequence 33073, A
44	68.5	10.8	238	4 US-10-425-115-246803	Sequence 246803,
45	68.5	10.8	365	5 US-10-732-923-10307	Sequence 10307, A

ALIGNMENTS

```
RESULT 1
US-10-617-038-30
; Sequence 30, Application US/10617038
; Publication No. US20040057963A1
GENERAL INFORMATION:
; APPLICANT: Andersen, Peter
; APPLICANT: Rosenkrands, Ida
; APPLICANT: Strym, Anette
; TITLE OF INVENTION: Therapeutic TB Vaccine
; FILE REFERENCE: SSI5AUSA
; CURRENT APPLICATION NUMBER: US/10/617,038
; PRIOR FILING DATE: 2003-07-11
; PRIOR APPLICATION NUMBER: DK PA 2002 01098
; PRIOR FILING DATE: 2002-07-13
; PRIOR APPLICATION NUMBER: US 60/401,725
; PRIOR FILING DATE: 2002-08-07
; NUMBER OF SEQ ID NOS: 187
; SOFTWARE: PatentIn version 3.2
SEQ ID NO 30
; LENGTH: 120
; TYPE: PR
; ORGANISM: Mycobacterium tuberculosis
US-10-617-038-30

Query Match      100.0%; Score 637; DB 4; Length 120;
Best Local Similarity 100.0%; Pred. NO. 1.4e-63;
Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MSTOPRRHSGIRAVGPYMAWAGRCGRIGRWGQAEAMNLAIWPRKVSATITYOTDRSH 60
      |||||||
DB      1 MSTOPRRHSGIRAVGPYMAWAGRCGRIGRWGQAEAMNLAIWPRKVSATITYOTDRSH 60
      |||||||

QY      61 DGRTRAVPDEITTSVSGWLSLGTQSLADELARAVRIGDMPAAVAIGEHLSVEIAVAV 120
      |||||||
DB      61 DGRTRAVPDEITTSVSGWLSLGTQSLADELARAVRIGDMPAAVAIGEHLSVEIAVAV 120
      |||||||

RESULT 2
US-10-425-114-39387
; Sequence 39387, Application US/10425114
; Publication No. US20040034888A1
GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
```

```
FILE REFERENCE: 38-21(53313)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 39387
LENGTH: 551
TYPE: PRT
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: 700089702_FLI.pcp
US-10-425-114-39387
```

```
Query Match      12.6%; Score 80; DB 4; Length 551;
Best Local Similarity 27.0%; Pred. No. 4.7;
Matches 43; Conservative 15; Mismatches 37; Indels 64; Gaps 10;
```

```
QY 6 PPH-----SGIRAVPYAMAGRC-----RIGRWGVHOEAMNLAIMPRKY-----47
DB 71 PPHLPGRGPRRRRRVAPQPHL-HGRGLREVVRAGEPVNH-LIHL-----PRRVPLPR 123
QY 48 -----QSATY-----QVTRSHDGRTRARVGDITSTVSGWLSLGTQSP 88
DB 124 PLGGHQLLRBAGLVGMPGHHHQGCTREHVSGSRGLPRPR-----QSGSQVP 173
QY 89 LADELARAVRIGDMPAIVIG-----HLSVEIAYA 119
DB 174 OGAMHARRARVQ--PALVELGETEVAAPRRHGVVEHVA 210
```

```
RESULT 3
US-10-794-514A-328
Sequence 328, Application US/10794514A
Publication No. US2005011234A1
GENERAL INFORMATION:
APPLICANT: Gradis, Thomas
APPLICANT: Laue, Reiner
APPLICANT: Diegel, Michael
APPLICANT: Vidovic, Damir
TITLE OF INVENTION: Compositions and Methods Employing Alternative
TITLE OF INVENTION: Reading Frame Polypeptides for the Treatment of
FILE REFERENCE: 11311.1003V
CURRENT APPLICATION NUMBER: US/10/794,514A
CURRENT FILING DATE: 2004-03-05
NUMBER OF SEQ ID NOS: 733
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 328
LENGTH: 1167
TYPE: PRT
ORGANISM: Derived From Human Sequence
US-10-794-514A-328
```

```
Query Match      12.5%; Score 79.5; DB 5; Length 1167;
Best Local Similarity 21.6%; Pred. No. 13;
Matches 40; Conservative 10; Mismatches 50; Indels 85; Gaps 8;
```

```
QY 4 QPRHSGR-AVGPYAMAGRCRIGRWGVHOE-----AMNLIATMH-----43
DB 310 QMHSQVSAASP---VEPCAMVMWMSVCERGQLVPISRLAARSLGAMHFCRAL 366
QY 44 -----PRKVQSAT-----IYVTRSHDGRTRARV-----PGDEI-----TS 74
DB 367 MCTOPTLPRSSQSSSKCLRWKRSQVITYTSGHRTACTLSASRTCKSGDEFCMAPTR 426
QY 75 TVSGW-----LSLGTQSPPLADELARAVRIGD 101
DB 427 PCKGMASAGWGAHGMWAVDWPSTITPTASACTRCPTGTSFETKTKLCSTLPTGQRTSV 486
QY 102 WPAAY 106
DB 487 WARAW 491
```

```
RESULT 4
US-10-282-122A-68099
Sequence 68099, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Lianghu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haebelbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Twawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Foreyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
```

```
FILE REFERENCE: EUTTRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
```

```
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 68099
LENGTH: 456
TYPE: PRT
ORGANISM: Pseudomonas putida
US-10-282-122A-68099
```

```
Query Match      12.2%; Score 77.5; DB 4; Length 456;
Best Local Similarity 23.3%; Pred. No. 7.2;
Matches 30; Conservative 15; Mismatches 43; Indels 41; Gaps 6;
```

```
QY 19 WAGRGGRGRGVHOEAMM-----NLAIWHPRVQSATIYQV-----56
DB 20 WAGQCGEGANNVTIVETIISGGAAPSSGAGWTGQQAQIKTLAEVSVVEIPRTTST 79
QY 57 --DR-----SHDGRTRARVGDITSTVSGWLSLGTQSPPLADELARAVRIGDMPAIVIG 109
DB 80 ELDRVLGGGLVDGVSIVLIGD-----PGICKSTILQTLQCN-IAVG-MPALVYTG 127
QY 110 EHLSEIAYV 118
DB 128 EESQOQVAM 136
```

```
RESULT 5
US-10-424-599-217918
Sequence 217918, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J
```



```
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 149941
; LENGTH: 279
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(279)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_50224C.1.pep
; US-10-437-963-149941

Query Match      11.7%; Score 74.5; DB 4; Length 279;
Best Local Similarity 29.0%; Pred. No. 8.7;
Matches 40; Conservative 13; Mismatches 52; Indels 33; Gaps 8;

QY      3  TORPRHSGIRAVGPVAMAGRCGR--IGRGVHOEAMNMLAIWHPKVSATTYQVTDNRSH 60
DB      120  TGRPRRGDEARPPPTAVTGAGSGGGRKAAAHGRSSGVAGARERTKAAT-----TTRE 175

QY      61  DERTA-----RVGDEI-----TSTVGWLSLSTGOSPLADELABAVRIGDW-PAAY- 106
DB      176  EGRTSGGDTXRRGGEEELGMEKATATTATTKVAGMRRRGMDSPP-----ABEGDVPVGR 227

QY      107  ---AIGE--HLSEIYVA 119
DB      228  XREAVAMAHMTAAELAA 245

RESULT 9
US-10-156-761-8113
; Sequence 8113, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 8113
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
; US-10-156-761-8113

Query Match      11.6%; Score 74; DB 4; Length 284;
Best Local Similarity 25.0%; Pred. No. 10;
Matches 17; Conservative 12; Mismatches 35; Indels 4; Gaps 1;

QY      36  MNMLAIWHPKVSATTYQVTDNRSHDGRARVPGDEITSTVGWLSLSTGOSPLADELAR 95
DB      88  VVNAGGLHAGLAARLRELIARHGHLKRVSHIRGDDVPSLDS----LQAGHCLPHLTS 143

QY      96  AVRIGDW 103
DB      144  GQPLSDWP 151

RESULT 10
US-10-282-122A-62103
; Sequence 62103, Application US/10282122A
```

```
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haeselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyekind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Foreyn, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.03A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See file wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 62103
; LENGTH: 535
; TYPE: PRT
; ORGANISM: Mycobacterium avium
; US-10-282-122A-62103

Query Match      11.6%; Score 74; DB 4; Length 535;
Best Local Similarity 28.2%; Pred. No. 22;
Matches 35; Conservative 15; Mismatches 46; Indels 28; Gaps 7;

QY      10  GIRAVGPVAMAGRCGRIGRWGVHOEAMNMLAIWHPKVSATTYQVTDNRSHDGRARVPG 69
DB      379  GIRVPGFITWPG-FARQSGDIGTAFSTVMDIA-----FTLELACTSHGTAAYR--G 426

QY      70  DEIT-----STVGWLSLSTGOSPLAD-----EL--ARAVRIGDW-----PAAYAIGEHLIS 113
DB      427  REVAPMRGRSLVGYLSGAADTVHDADTETGWELEFERRAIRQGNWKAHLIPAVYGGSWQL 486

QY      114  VEIA 117
DB      487  YDLA 490

RESULT 11
US-09-864-761-43789
; Sequence 43789, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
```

```

; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aecm1ca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 43789
; LENGTH: 69
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC005920.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.68
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.68
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.53
; OTHER INFORMATION: EXPRESSED IN PETUL LIVER, SIGNAL = 0.69
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.61
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.97
; OTHER INFORMATION: SWISSPROT HIT: P17691, EVALU8 8.80e-01
; OTHER INFORMATION: EST_HUMAN HIT: A0143722.1, EVALU8 6.00e-25
US-09-864-761-43789

Query Match 11.5%; Score 73; DB 3; Length 69;
Best Local Similarity 28.1%; Pred. No. 2.4;
Matches 18; Conservative 15; Mismatches 27; Indels 4; Gaps 1;

QY 54 QVTRSHDGRTRVPEDEITSTVSGWLSLGTQSPDLADLARAARI-GDMPAAVAIGELHS 113
DB 2 ELSDVSGGSKATP---ASTANSDVATIPDTPIKENEKGVKVTADAPNKSLSKSHTE 57
QY 114 VEIA 117
DB 58 VOYA 61

RESULT 12
```

```

US-10-106-698-6371
; Sequence 6371, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 6371
; LENGTH: 258
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (140)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-106-698-6371

Query Match 11.5%; Score 73; DB 4; Length 258;
Best Local Similarity 27.0%; Pred. No. 12;
Matches 37; Conservative 14; Mismatches 40; Indels 46; Gaps 9;

QY 9 SGIRVGPY---AMAGRCGRIG-----RMGVOBAMMNLAIW-HPRKY 47
DB 121 STVREVGHSYSCSMTCNGCAXLAALDEPGYAQINTEADKGLIDBEAL---VWHSRKG 176
QY 48 QSATYQVTRSHDGRTRVPEDEITSTVSGWL---SLGTQ--SPLDE-----LARA 96
DB 177 KIITRAQVSDRPNKG-----AIYMTYQWVIGACNELVTEMLSPITTPRYKCAVAV 228
QY 97 VAIQDMPAA--YALGEH 111
DB 229 EPIADQRAAEQYVIDEY 245

RESULT 13
US-10-732-923-10308
; Sequence 10308, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgeton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 10308
; LENGTH: 365
; TYPE: PRT
; ORGANISM: Escherichia coli CFT073
US-10-732-923-10308

Query Match 11.5%; Score 73; DB 5; Length 365;
Best Local Similarity 26.0%; Pred. No. 18;
Matches 27; Conservative 17; Mismatches 32; Indels 28; Gaps 6;

QY 19 WAGRCG-----RGRW-----GVHOE-AMMNLAIWHPKVOGATYQVTRSHDGRTRV 67
DB 69 YAGPFIHSQECIAAMKKITAGVHAENGHMAVOLMHTGRISHASIQ----- 114
QY 68 PGDEITSTVSGWLSLGTQSPDLADLARAARI-GDMPAAVAIGE 110
```

Db 115 PGQ-APVAPSALSAGTRTSLRDENQAIRVETSMPRALTEGE 156

RESULT 14
US-10-104-047-2749
; Sequence 2749, Application US/10104047
; Publication No. US20030236392A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20030236392A1e1 full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2749
; LENGTH: 449
; TYPE: PR1
; ORGANISM: Homo sapiens
; US-10-104-047-2749

```

Query Match Similarity      11.5%   Score 73; DB 4; Length 449;
Best Local Similarity     28.1%   Pred. No. 23;
Matches      18; Conservative    15; Mismatches    27; Indels      4; Gaps      1

OY          54 QVTRSDHGRFARVPGEIDITSTVSCWLSLGTQSPFLADELARVRIGDMPAAVAIGEHLS 113
           ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB         249 ELDSVSOQGSGSKATTP-----ASTANSDVATIPTDTPLKEENEGFVKVTDAFPNKSISKHTE 304
OY          114 VEIA 117
           ||::|
DB         305 VOVA 308

RESULT 15
US-09-815-242-10450
; Sequence 10450, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zykkind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: EUTRA 011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10450
LENGTH: 715
TYPE: PRT
```

```

; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(715)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-815-242-10450

Query Match          11.5%; Score 73; DB 3; Length 715;
Best Local Similarity 27.0%; Pred. No. 40;
Matches 37; Conservative 14; Mismatches 40; Indels 46; Gaps 9

Qy      9  SGIRAVGPY---AMAGRCGRIG-----RMGYHQAEMNNLAIW-HPRKY 47
      | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      578 STVREYGHVSCRSMTGNCALALADLBEGVAQINTEDAKRIGIEDAL-----VWVHSRK 633
      | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy      48 QASNTTYQVTDNRSHDGRITARVDEDELTSTVSSWL---SELGTQ--SPLADE-----LAAR 96
      | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      634 KIITRAQVSDRPNKG-----ALYMTYQWMIAGCNELVTENLSPITKPEYKCAVRR 685
      | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy      97 VRIGDWPAA--YAIGEH 111
      | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      686 EPIADQRAALAEQYVIDEY 702
      | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

Search completed: March 23, 2006, 06:28:20
Job time : 26.2352 secs

```

Query Match      11.5%; Score 73; DB 3; Length 715;
Best Local Similarity 27.0%; Pred. No. 40;
Matches 37; Conservative 14; Mismatches 40; Indels 46; Gaps 9

Qy      9  SGIRAVGPY---AMAGRCGRIG-----RQVHQEAMNLAIW-HPKV 47
      | : | | | : | | | : | : | : | : | : | : | : | : |
Db      578 STVREVGHYSCRSMTGNCALALADBEQYQINTEDAKRIEDEL---VMVHSRKG 633
Qy      48  QSATIVQVTDNRSHDGRRTARVPODEITSTVSGWL---SELGTQ--SPLADE-----LARA 96
      | : | | | | | | | : | | | : | : | : | : | : | : |
Db      634 KIITRAQVSDRENKG-----ATYMTYQWIGACNELVTENLSPITTKPEYKCAVRY 685
Qy      97  VRIGDWPAA--YAIGEY 111
      | | | | | | | : |
Db      686 EPIADQRAAEQYVIDEY 702

Search completed: March 23, 2006, 06:28:20
Job time : 26.2352 secs

```

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd

OM protein - protein search, using SW model

Run on: March 23, 2006, 06:17:23 ; Search time 2.8797 Seconds
(without alignments)
1192.746 Million cell updates/sec

Title: US-10-617-038-30

Sequence: 1 MSTQRPRHSGIRAVGPYAWA.....DWPAYAIGEHL SVEIAVAV 120

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 169630 seqs, 28622889 residues

Total number of hits satisfying chosen parameters: 169630

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

```
Post-processing:  Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries
```

```
Database : Published Applications AA.NWS:*
1: /cgn2_6/prodataa/1/pubpaal/US08_NEM_PUB.dep:*
2: /cgn2_6/prodataa/1/pubpaal/US06_NEM_PUB.dep:*
3: /cgn2_6/prodataa/1/pubpaal/US07_NEM_PUB.dep:*
4: /cgn2_6/prodataa/1/pubpaal/PCT_NEM_PUB.dep:*
5: /cgn2_6/prodataa/1/pubpaal/US05_NEM_PUB.dep:*
6: /cgn2_6/prodataa/1/pubpaal/US10_NEM_PUB.dep:*
7: /cgn2_6/prodataa/1/pubpaal/US11_NEM_PUB.dep:*
8: /cgn2_6/prodataa/1/pubpaal/US60_NEM_PUB.dep:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	80.5	12.6	520	7	US-11-087-099-5634	Sequence 5634, App
2	73	11.5	449	7	US-11-072-512-2749	Sequence 2749, App
3	71.5	11.2	759	6	US-10-467-657-7722	Sequence 2722, App
4	69.5	10.9	472	6	US-10-509-464-5	Sequence 5, App11
5	68.5	10.8	446	7	US-11-146-428-47	Sequence 47, App1
6	68	10.7	571	7	US-11-072-512-3814	Sequence 3814, App
7	66.5	10.4	231	7	US-11-000-463-745	Sequence 745, App
8	66.5	10.4	486	7	US-11-000-463-773	Sequence 773, App
9	66	10.4	491	7	US-11-031-206-72	Sequence 72, App1
10	64	10.0	1627	6	US-10-821-234-1283	Sequence 1283, App
11	63.5	10.0	410	7	US-11-049-536-352	Sequence 352, App1
12	63.5	10.0	108	7	US-11-096-568A-18440	Sequence 18440, App
13	63.5	10.0	412	7	US-11-096-568A-18439	Sequence 18439, App
14	63.5	10.0	417	7	US-11-096-568A-18438	Sequence 18438, App
15	63.5	10.0	447	7	US-11-096-568A-24503	Sequence 24503, App
16	63.5	10.0	510	7	US-11-096-568A-24502	Sequence 24502, App
17	63	9.9	224	7	US-11-072-512-2601	Sequence 2601, App
18	63	9.9	328	6	US-10-517-939-0	Sequence 20, App1
19	62.5	9.8	328	7	US-11-055-822-554	Sequence 554, App1
20	62.5	9.8	328	7	US-11-055-822-556	Sequence 556, App1
21	62.5	9.8	341	7	US-11-087-099-2030	Sequence 2030, App
22	62.5	9.8	432	6	US-10-467-657-1890	Sequence 1890, App
23	62.5	9.8	432	6	US-10-467-657-6128	Sequence 6128, App
24	62.5	9.8	447	7	US-11-087-099-5814	Sequence 5814, App
25	62	9.7	254	7	US-11-054-515-1854	Sequence 1854, App

26	62	9.7	254	7	US-11-054-515-1839	Sequence 1857, App
27	61.5	9.7	127	7	US-11-072-512-9257	Sequence 2939, App
28	61.5	9.7	253	7	US-11-054-515-1249	Sequence 1249, App
29	61.5	9.7	283	6	US-10-653-150-50	Sequence 50, App1
30	61.5	9.7	283	7	US-11-073-457-50	Sequence 50, App1
31	61.5	9.7	283	7	US-11-073-460-50	Sequence 50, App1
32	61.5	9.7	319	7	US-11-096-5686A-17724	Sequence 17724, App
33	61.5	9.7	341	7	US-11-087-039-9325	Sequence 9325, App
34	61.5	9.7	358	7	US-11-046-428-112	Sequence 112, App
35	61	9.6	289	7	US-11-096-5686A-23987	Sequence 23987, App
36	61	9.6	324	7	US-11-096-5686A-23986	Sequence 23986, App
37	61	9.6	356	7	US-11-087-039-1946	Sequence 1946, App
38	61	9.6	424	6	US-10-467-657-90	Sequence 90, App1
39	61	9.6	424	6	US-10-467-657-90	Sequence 90, App1
40	60.5	9.5	137	6	US-10-475-075-196	Sequence 196, App
41	60.5	9.5	137	6	US-10-475-075-479	Sequence 479, App
42	60.5	9.5	398	7	US-11-053-5544A-252	Sequence 252, App
43	60	9.4	520	7	US-11-055-832-862	Sequence 862, App
44	60	9.4	615	6	US-10-995-561-940	Sequence 940, App
45	59.5	9.3	218	7	US-11-096-5686A-19823	Sequence 19823, App

ALIGNMENTS

```

RESULT 1
US-11-087-099-5634
: Sequence 5634, Application US/11087099
: Publication No. US20060041961A1
: GENERAL INFORMATION:
: APPLICANT: Abad, Mark S. et al.
: TITLE OF INVENTION: Genes and Uses for Plant Improvement
: FILE REFERENCE: 38-21(53450)B EP
: CURRENT APPLICATION NUMBER: US/11/087,099
: CURRENT FILING DATE: 2005-03-22
: NUMBER OF SEQ ID NOS: 12464
: SEQ ID NO 5634
: LENGTH: 520
: TYPE: PR1
: ORGANISM: Agrobacterium tumefaciens str. C58 (Cereon)
US-11-087-099-5634

Query Match 12.6%; Score 80.5; DB 7; Length 520;
Best Local Similarity 28.0%; Pred. No. 0.32;
Matches 28; Conservative 9; Mismatches 56; Indels 7; Gaps 2

Oy 2 STQPRRISGRVANGPYAMAGRCGRIGMGVGHQEMMMLAIWHPRKVSATITYQVTRSD 61
Db 8 SPEHPFHS--AGNNQALAAARVYMASWNISNPMSASAIWLPERISPMITIDIKNIAMD 65
Oy 62 GRTARVPGDEITSTVSGMLSELGTQSPLEADLARAVRIGD 101
Db 66 RNAVLHPFTQLKDPASGKLG-----PTVETGGRIRID 100

RESULT 2
US-11-072-512-2749
: Sequence 2749, Application US/11072512
: Publication No. US20060028945A1
: GENERAL INFORMATION:
: APPLICANT: ISOGAI, TAKAO
: APPLICANT: SUGIYAMA, TOMOYASU
: APPLICANT: OTSUKI, TETSUJI
: APPLICANT: WAKAMATSU, AI
: APPLICANT: SATO, HIROYUKI
: APPLICANT: ISHII, SHIZUKO
: APPLICANT: YAMAMOTO, JUN-ICHI
: APPLICANT: ISONO, YUUKO
: APPLICANT: HIO, YURI
: APPLICANT: OTSUKA, KAORU
: APPLICANT: NAGAI, KEIICHI
: APPLICANT: IRIE, RYOTARO
: APPLICANT: TAMECHIKA, ICHIRO

```

```

; APPLICANT: SEKI, NAOHITO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOYUKI
; APPLICANT: NAGAHARA, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: 084335-0191
; CURRENT APPLICATION NUMBER: US/11/072,512
; CURRENT FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: US 60/350,978
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: JP 2001-379298
; PRIOR FILING DATE: 2001-11-05
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2749
; LENGTH: 449
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-072-512-2749

```

Query Match 11.5%; Score 73; DB 7; Length 449;

Best Local Similarity 28.1%; Pred. No. 1.9; Mismatches 27; Indels 4; Gaps 1;

Matches 18; Conservative 15; Mismatches 27; Indels 4; Gaps 1;

```

QY 54 QVTRSHDGRARVPQDEITSTVSGWLSLGTQSPDLDELARAVRIGDMPAAVAIGEHLIS 113
DB 249 ELSDVSGGSKATTP---ASTANSDVATIPDTPLKENEGFVKVTDAPNKSEISKHLE 304
QY 114 VEIA 117
DB 305 VQVA 308

```

RESULT 3

```

US-10-467-657-2722
; Sequence 2722, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; APPLICANT: PONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIIGNANT Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqMin99, version 1.04
; SEQ ID NO 2722
; LENGTH: 759
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-2722

```

Query Match 11.2%; Score 71.5; DB 6; Length 759;

Best Local Similarity 25.2%; Pred. No. 5.2; Mismatches 53; Indels 13; Gaps 3;

```

QY 6 PRHSGRIRAVGPYAMGRCGRIGRWGVHDEAMNMLAIWHPRKV-----QSATTYQ 54
DB 309 PQRSVLMALAPAMWRRLSNAMATWQALAAVLLFDPLATLVGCTWMLSPGLVAALITIA 368
QY 55 VTRSHDGR-RTARVPQDEITSTVS-GWLSLGTQSPDLDELARAVRI 99
DB 369 CAGCLYEGKRQTAIVAGQWMAVSILVILGYLPLASLPLVPLVNAVSI 415

```

RESULT 4

US-10-509-464-5

```

; Sequence 5, Application US/10509464
; Publication No. US20050244826A1
; GENERAL INFORMATION:
; APPLICANT: NIEHRS, CHRISTOP
; APPLICANT: MAO, BINGYU
; TITLE OF INVENTION: ASSOCIATIONS FOR DIAGNOSIS AND THERAPY OF DISEASES
; TITLE OF INVENTION: ASSOCIATED WITH ABERRANT EXPRESSION OF KEMEN AND/OR
; FILE REFERENCE: WNT
; CURRENT APPLICATION NUMBER: US/10/509,464
; CURRENT FILING DATE: 2004-09-27
; PRIOR APPLICATION NUMBER: PCT/EP03/03277
; PRIOR FILING DATE: 2003-03-28
; PRIOR APPLICATION NUMBER: EP 02 008 650.0
; PRIOR FILING DATE: 2002-04-17
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patent In Ver. 3.3
; SEQ ID NO 5
; LENGTH: 473
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-509-464-5

```

Query Match 10.9%; Score 69.5; DB 6; Length 473;

Best Local Similarity 27.8%; Pred. No. 5; Mismatches 44; Indels 21; Gaps 4;

Matches 30; Conservative 13; Mismatches 44; Indels 21; Gaps 4;

```

QY 3 TORPHSGIRAVGPYAMGRCG---RIGRWGVHDEAMNMLAIWHPRKVQSATTYQVTR 58
DB 361 TPSPHPQOTAPGSHSMAPSVGANSRVEGTVV--GLATLLILTVTAIVAKILHVTFK 418
QY 59 SH-----DGRARVPD-----EITSTVSGWLSLGTQSPDLAD 91
DB 419 SHRVPASGDLRDCRPGASGDITWTFYBPSTTISIFKKKLGQSQODD 466

```

RESULT 5

```

US-11-146-428-47
; Sequence 47, Application US/11146428
; Publication No. US20060051847A1
; GENERAL INFORMATION:
; APPLICANT: GUNNARSSON, NINA KATARINA
; APPLICANT: FORSTER, JOCHEN
; APPLICANT: NEILSEN, JENS BREDEL
; TITLE OF INVENTION: METABOLICALLY ENGINEERED CELLS FOR THE
; FILE REFERENCE: 671306-2001.1
; CURRENT APPLICATION NUMBER: US/11/146,428
; CURRENT FILING DATE: 2005-06-06
; PRIOR APPLICATION NUMBER: 60/577,245
; PRIOR FILING DATE: 2004-06-04
; NUMBER OF SEQ ID NOS: 227
; SOFTWARE: Patent In Ver. 3.3
; SEQ ID NO 47
; LENGTH: 446
; TYPE: PRT
; ORGANISM: Cryplococcus curvatus
US-11-146-428-47

```

Query Match 10.8%; Score 68.5; DB 7; Length 446;

Best Local Similarity 21.2%; Pred. No. 6; Mismatches 44; Indels 29; Gaps 4;

```

QY 20 AGRCGRIGRWGVHQ--EAMNMLAIW-----HPRKVQSATTI----- 52
DB 97 AHTAKILMFTTYQVLAVGVTGIVNVAHEGCHQAYSSKTIINNAVGVVHLSILVPHYS 156
QY 53 QVTRSHDGRARVPQDEITSTVSGWLSLGTQSPDLDELARAVRIGDMPAA 105
DB 157 WRISHGRHATYGHLTRDEV--FVPRTRREQIGIQAPKTEEEKKGINVAMPQQA 207

```

RESULT 6

```
US-11-072-512-3814
; Sequence 3814, Application US//11072512
; Publication No. US20060029945A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUYI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUOKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEITICH
; APPLICANT: IRIE, RYO太郎
; APPLICANT: TIMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHKO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOTYUKI
; APPLICANT: NAGAARI, KENJI
; APPLICANT: MASUDO, YASUHICO
; TITLE OF INVENTION: Novel full length cDNA
FILE REFERENCE: 084335-0191
CURRENT APPLICATION NUMBER: US/11/072,512
CURRENT FILING DATE: 2005-03-07
PRIOR APPLICATION NUMBER: US 60/350,978
PRIOR FILING DATE: 2002-01-25
PRIOR APPLICATION NUMBER: JP 2001-379298
PRIOR FILING DATE: 2001-11-05
NUMBER OF SEQ ID NOS: 4096
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3814
LENGTH: 571
TYPE: PRT
ORGANISM: Homo sapiens
US-11-072-512-3814

Query Match      10.7% Score 68; DB 7; Length 571;
Best Local Similarity 28.3%; Pred.No.9.2;
Matches 30; Conservative 15; Mismatches 29; Indels 32; Gaps 7

Qy          Db    29   GCWNO---EAMMNLAIHPKRVQSATTIQVTDNRSHDGTARVPDEDTIS--TVSGMWLSE 82
           |||::||::||::||::||::||::||::||::||::||::||::||::||::||::||:
Db         4     WGGERRGAGEGSPKLAVAYATRKTRR--VRSDGDQWTLG----YPDDOWSSGFPSYSWMKN 57

Qy          Db    83 LGTOS-----PLAD-----ELRAVRLIGMPMAVAIGEHL 112
           ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db         58 VGSERKGEGALDPQHVDIVREDLSGLHRPARSGDVPGV---EHV 99

RESULT 7
US-11-000-463-745
; Sequence 745, Application US//11000463
; Publication No. US20050266423A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Chen, Rui-hong
; APPLICANT: Qian, Xiaohong B.
; APPLICANT: Wang, Zhiwei
; APPLICANT: Wehrman, Tom
; APPLICANT: Zhang, Jie
; APPLICANT: Zhou, Ping
; APPLICANT: Cao, Yi-Cheng
; APPLICANT: Dimaenac, Radjoje T.
TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
FILE REFERENCE: 785CIPACN
CURRENT APPLICATION NUMBER: US/11/000,463
CURRENT FILING DATE: 2004-11-29
PRIOR APPLICATION NUMBER: 2004/11,265
PRIOR FILING DATE: 2002-11-08
```

```

PRIORITY APPLICATION NUMBER: PCT/US01/02623
PRIOR FILING DATE: 2001-01-25
PRIOR APPLICATION NUMBER: 09/922,279
PRIOR FILING DATE: 2001-08-03
PRIOR APPLICATION NUMBER: 09/491,404
PRIOR FILING DATE: 2000-01-25
PRIOR APPLICATION NUMBER: 09/611,746
PRIOR FILING DATE: 2000-07-17
PRIOR APPLICATION NUMBER: 09/631,451
PRIOR FILING DATE: 2000-09-15
SOFTWARE: PasteSeq for Windows Version 3.0
SEQ ID NO 745
LENGTH: 231
TYPE: PRT
ORGANISM: Homo sapiens
US-11-000-463-745

Query Match      10.4%; Score 66.5; DB 7; Length 231;
Best Local Similarity 24.4%; Pred. No. 4.5;
Matches 21; Conservative 12; Mismatches 34; Indels 19; Gaps 3

Cy      18 AMARCGRIGWGHQAEAM-----NLATWPRKVSATIQVTDHRSIDGRTARVPGEI 72
Db      17 AMPALCGYKKRRKHGVFLTLTTSVAIW---VVMITWTYGKNHSPTWDPLAI 72
          ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Oy      73 TSTVSGW-----LSELGTOSP 88
Db      73 ALAANAFAVLFYIPIPSVQTAKSP 98
          : | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 8
US-11-000-463-273
Sequence 273, Application US//11000463
Publication No. US20050266423A1
GENERAL INFORMATION:
APPLICANT: Tang, Y Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Chen, Rui-hong
APPLICANT: Qian, Xiaohong B.
APPLICANT: Wang, Zhiwei
APPLICANT: Wehrman, Tom
APPLICANT: Zhang, Jie
APPLICANT: Zhou, Ping
APPLICANT: Cao, Yi-Cheng
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
FILE REFERENCE: 785CIP4CN
CURRENT APPLICATION NUMBER: US//11/000,463
CURRENT FILING DATE: 2004-11-29
PRIOR APPLICATION NUMBER: 10/291,265
PRIOR FILING DATE: 2002-11-08
PRIOR APPLICATION NUMBER: PCT/US01/02623
PRIOR FILING DATE: 2001-01-25
PRIOR APPLICATION NUMBER: 09/922,279
PRIOR FILING DATE: 2001-08-03
PRIOR APPLICATION NUMBER: 09/491,404
PRIOR FILING DATE: 2000-01-25
PRIOR APPLICATION NUMBER: 09/611,746
PRIOR FILING DATE: 2000-07-17
PRIOR APPLICATION NUMBER: 09/631,451
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: 09/633,870
PRIOR FILING DATE: 2000-09-15
NUMBER OF SEQ ID NOS: 944
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 273
LENGTH: 466
TYPE: PRT
ORGANISM: Homo sapiens

```


Best Local Similarity 28.2%; Pred. No. 4;
Matches 24; Conservative 7; Mismatches 29; Indels 25; Gaps 4;

Qy 38 NLAIW---HPRKQSAITVOTDR-----SHDGTARVPGEITVSGWLSLG 84
Db 34 NYVSWYQGHGKAPKLLITIDVSPBGVSNRPSGKSGNTASL-----TISGLAE-- 84

Qy 85 TQSLADELARAVRIGDWPAAVAIG 109
Db 85 ---DEADYCGSRVTSVRSYVFG 106

RESULT 12

US-11-096-568A-18440
; Sequence 18440, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:

; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 18440
; LENGTH: 408
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays

; NAME/KEY: misc_feature
; LOCATION: (1)..(408)
; OTHER INFORMATION: Ceres Seq. ID no. 12365194
US-11-096-568A-18440

Query Match 10.0%; Score 63.5; DB 7; Length 408;
Best Local Similarity 47.5%; Pred. No. 20;
Matches 19; Conservative 3; Mismatches 9; Indels 9; Gaps 4;

Qy 12 RAVGPYAMAGRCGRIGR---WGVHQBAMNMLAIWHPKV 47
Db 263 RAAGA---VGRCCR-GRLAVAGVHQPRL-LPRMHPRV 297

Query Match 10.0%; Score 63.5; DB 7; Length 408;
Best Local Similarity 47.5%; Pred. No. 20;
Matches 19; Conservative 3; Mismatches 9; Indels 9; Gaps 4;

Qy 12 RAVGPYAMAGRCGRIGR---WGVHQBAMNMLAIWHPKV 47
Db 263 RAAGA---VGRCCR-GRLAVAGVHQPRL-LPRMHPRV 297

RESULT 13

US-11-096-568A-18439
; Sequence 18439, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:

; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 18439
; LENGTH: 412
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays

; NAME/KEY: misc_feature
; LOCATION: (1)..(412)
; OTHER INFORMATION: Ceres Seq. ID no. 12365193
US-11-096-568A-18439

Query Match 10.0%; Score 63.5; DB 7; Length 412;
Best Local Similarity 47.5%; Pred. No. 20;
Matches 19; Conservative 3; Mismatches 9; Indels 9; Gaps 4;

Qy 12 RAVGPYAMAGRCGRIGR---WGVHQBAMNMLAIWHPKV 47
Db 267 RAAGA---VGRCCR-GRLAVAGVHQPRL-LPRMHPRV 301

RESULT 14

US-11-096-568A-18438
; Sequence 18438, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:

; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 18438
; LENGTH: 417
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays

; NAME/KEY: misc_feature
; LOCATION: (1)..(417)
; OTHER INFORMATION: Ceres Seq. ID no. 12365192
US-11-096-568A-18438

Query Match 10.0%; Score 63.5; DB 7; Length 417;
Best Local Similarity 47.5%; Pred. No. 20;
Matches 19; Conservative 3; Mismatches 9; Indels 9; Gaps 4;

Qy 12 RAVGPYAMAGRCGRIGR---WGVHQBAMNMLAIWHPKV 47
Db 272 RAAGA---VGRCCR-GRLAVAGVHQPRL-LPRMHPRV 306

Query Match 10.0%; Score 63.5; DB 7; Length 417;
Best Local Similarity 47.5%; Pred. No. 20;
Matches 19; Conservative 3; Mismatches 9; Indels 9; Gaps 4;

Qy 12 RAVGPYAMAGRCGRIGR---WGVHQBAMNMLAIWHPKV 47
Db 272 RAAGA---VGRCCR-GRLAVAGVHQPRL-LPRMHPRV 306

RESULT 15

US-11-096-568A-24503
; Sequence 24503, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:

; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 24503
; LENGTH: 447
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays

; NAME/KEY: misc_feature
; LOCATION: (1)..(447)
; OTHER INFORMATION: Ceres Seq. ID no. 12435494
; FEATURE: NAME/KEY: misc_feature
; LOCATION: (52)..(52)
; OTHER INFORMATION: Xaa is any aa, unknown or other
; FEATURE: NAME/KEY: misc_feature
; LOCATION: (58)..(58)
; OTHER INFORMATION: Xaa is any aa, unknown or other
; FEATURE: NAME/KEY: misc_feature
; LOCATION: (59)..(59)
; OTHER INFORMATION: Xaa is any aa, unknown or other
US-11-096-568A-24503

Query Match 10.0%; Score 63.5; DB 7; Length 447;
Best Local Similarity 25.0%; Pred. No. 22;
Matches 23; Conservative 7; Mismatches 41; Indels 21; Gaps 3;

Qy 11 IRVGPYAMAGRCGRIGWGVHQBAMNMLAIWHPKV 70
Db 4 LRVVG---WMTTAAAPBEMGVHRCRCGTPLARBPAAVTSU-----GAGQ 45

Query Match 10.0%; Score 63.5; DB 7; Length 447;
Best Local Similarity 25.0%; Pred. No. 22;
Matches 23; Conservative 7; Mismatches 41; Indels 21; Gaps 3;

Qy 11 IRVGPYAMAGRCGRIGWGVHQBAMNMLAIWHPKV 70
Db 4 LRVVG---WMTTAAAPBEMGVHRCRCGTPLARBPAAVTSU-----GAGQ 45

Query Match 10.0%; Score 63.5; DB 7; Length 447;
Best Local Similarity 25.0%; Pred. No. 22;
Matches 23; Conservative 7; Mismatches 41; Indels 21; Gaps 3;

Qy 11 IRVGPYAMAGRCGRIGWGVHQBAMNMLAIWHPKV 70
Db 4 LRVVG---WMTTAAAPBEMGVHRCRCGTPLARBPAAVTSU-----GAGQ 45

Oy 71 EITSTVSGWLSEL--GTOSPLADELARAARI 99
| | | | | : | | | | :
Db 46 LRVGTExGWLMDXGGGGAGADYAREMEVAARV 77
Search completed: March 23, 2006, 06:29:41
Job time : 3.8797 secs